National Rivers and Streams Assessment 2008–2009

Technical Report

DRAFT

U.S. Environmental Protection Agency
Office of Wetlands, Oceans and Watersheds
Office of Research and Development
Washington, DC 20460

February 26, 2013

Table of Contents

A National Rivers and Streams Assessment Survey Design: 2008–2009 5

A.1 Background information 5

A.2 Evaluation process 9

A.3 Statistical analysis 10

A.4 Literature cited... 10

B Reference Condition for the NRSA 2008–2009 11

B.1 Background information 11

B.2 Methodology... 11

B.3 Literature cited... 16

C Benthic Macroinvertebrate Assemblage for NRSA 2008–2009 17

C.1 Background information 17

C.2 Data preparation... 18

C.3 Multimetric Index development 19

C.4 O/E: predictive (RIVPACS) models 21

C.5 Literature cited... 23

D Fish Community Assemblage for NRSA 2008–2009.. 24

D.1 Background information 24

D.2 Methods 25

D.3 Fish Multimetric Index development... 26

D.4 FMMI performance 35

D.5 Thresholds for assigning ecological condition 40

D.6 Literature cited... 41

E Periphyton Assemblage Analysis for the NRSA 2008–2009 44

E.1 Background information 44

E.2 Methods 45

E.3 Results 50

E.4 Discussion 54

E.5 Literature cited 55

F Water Chemistry Analysis for NRSA 2008–2009 70

F.1 Background information 70

F.2 Threshold development 70

F.3 Literature cited 71

G Physical Habitat Assessment for the NRSA 2008–2009 72

G.1 Background information 72

G.2 Methods 73

G.3 Physical habitat condition indicators in the NRSA 76

G.4 Estimating reference condition for physical habitat 82

G.5 Response of the physical habitat indicators to human disturbance 83

G.6 Literature cited 84

Appendix G-1. Reference condition models 96

H Enterococci Indicator for the NRSA 2008–2009 118

H.1 Background information 118

H.2 Methods 118

H.3 Thresholds 118

H.4 Literature cited 119

I Human Health Fish Tissue Indicator — Mercury 121

I.1 Background 121

I.2 Field fish collection 121

I.3 Mercury analysis and human health screening values 122

I.4 Literature cited 122

J Relative Extent, Relative Risk and Attributable Risk for the NRSA 2008–2009 124

J.1 Background information 124

K Change Analysis from Wadeable Streams Assessment 2004 to National Rivers and Streams Assessment 2008–2009 125

K.1 Background information 125

K.2 Overall change analysis 125

K.3 Caveats for looking at change analysis 126

National Rivers and Streams Assessment Survey Design: 2008–2009

1.1 Background information

The design requirements for the National Rivers and Streams Assessment are to produce:

1. Estimates of the 2008–2009 status of flowing waters nationally and regionally (nine aggregated Omernik ecoregions).
2. Estimates of the 2008–2009 status of wadeable streams and non-wadeable rivers nationally and regionally (nine aggregated Omernik ecoregions).
3. Estimates of the 2008–2009 status or urban flowing waters nationally.
4. Estimates of the change in status in wadeable streams between 2008-2009 and 2004, nationally and regionally (nine aggregated Omernik ecoregions).

A secondary objective is to have each state sample approximately an equal number of sites (37–38).

A.1.1 *Target population*

The target populations consists of all streams and rivers within the 48 contiguous states that have flowing water during the study index period, excluding portions of tidal rivers up to head of salt. The study index period extends from April/May to September and is generally characterized by low flow conditions. The target population includes the Great Rivers. Run-of-the-river ponds and pools are included while reservoirs are excluded.

A.1.2 *Sample frame*

The sample frame was derived from the National Hydrography Dataset (NHD), in particular NHD-Plus. Attributes from NHD-Plus and additional attributes added to the sample frame that are used in the survey design include: (1) state, (2) EPA Region, (3) NAWQA Mega Region, (4) Omernik Ecoregion Level 3 (NACEC version), (4) WSA aggregated ecoregions (nine and three regions), (5) Strahler order, (6) Strahler order categories (1^{st}, 2^{nd}, …, 7^{th}, and 8^{th}+), (6) FCODE, (7) Urban, and (8) Frame07.

The version of NHD-Plus used includes two separate Strahler order calculations, one of which is included in the publicly available NHD-Plus version. The other Strahler order calculation (SO attribute name) more accurately reflects the true Strahler order and is used for the survey design. The StrahCat attribute collapses 8^{th}, 9^{th}, and 10^{th} order rivers into a single category.

The Urban attribute was created by intersecting a modified version of the Census Bureau national urban boundary GIS coverage with NHD-Plus. The Census Bureau's boundaries were buffered 100 meters to include a majority of stream features intersecting and coincident with urban areas. Where this buffer did not completely gather all the river features within the urban areas (rivers intersecting cities are excluded from the Census Bureau's urban areas), the NHD-Plus river area (polygon) features were clipped at a 3-kilometer buffer around the urban areas and combined with the buffered urban area to create the modified urban database. If a stream or river segment was within this boundary, it was designated as "Urban"; otherwise it was "NonUrban."

FCODE comes directly from NHD-Plus and was used to identify which segments in NHD were included in the sample frame. The attribute Frame07 identifies each segment as either "Include" or "Exclude." Frame07 was created so that segments included in the sample frame could be easily identified. FCODE values included in the GIS shapefile:

Included in FW08 sample frame (Frame07 = "Include"):

33400	Connector
33600	Canal/Ditch
42801	Pipeline: Pipeline Type = Aqueduct; Relationship to Surface = At or Near
46000	Stream/River
46003	Stream/River (Intermittent)
46006	Stream/River (Perennial)
58000	Artificial Path (removed from dataset if coded through Lake/Pond and Reservoirs)

Excluded in FW08 sample frame (Frame07 = "Exclude")

42800	Pipeline
42802	Pipeline: Pipeline Type = Aqueduct; Relationship to Surface = Elevated
42803	Pipeline: Pipeline Type = Aqueduct; Relationship to Surface = Underground
42804	Pipeline: Pipeline Type = Aqueduct; Relationship to Surface = Underwater
42806	Pipeline: Pipeline Type = General Case; Relationship to Surface = Elevated
4280	Pipeline: Pipeline Type = General Case; Relationship to Surface = Underground
42809	Pipeline: Pipeline Type = Penstock; Relationship to Surface = At or Near
42811	Pipeline: Pipeline Type = Penstock; Relationship to Surface = Underground
42813	Pipeline: Pipeline Type = Siphon
56600	Coastline

Rivers with a Strahler order greater than or equal to 5th order that had FCODE equal to 46003 (intermittent) were included in the FW08 sample frame for all states west of 96 degrees longitude (North Dakota to Texas and states west). This was done to ensure that all large rivers in the more arid west were included regardless of NHD-Plus intermittent code.

A.1.3 *Survey design*

The survey design consists of two major components in order to address the dual objectives of (1) estimating current status for all flowing waters and (2) estimating change in status for wadeable streams from the 2004 Wadeable Streams Assessment. These two components are NRSA design and WSA_Revisit design. A Generalized Random Tessellation Stratified (GRTS) survey design for a linear resource was used for the NRSA design and a GRTS survey design for a finite resource was used for the WSA_Revisit design. The design includes reverse hierarchical ordering of the selected sites.

A.1.4 *Stratification*

The survey design was explicitly stratified by state for the NRSA design. The original WSA design had several strata (EMAP West, New England, Virginia, Iowa, and remaining eastern states combined). The WSA_Revisit design ignored these strata in the selection of the subset of sites from the WSA to be revisited as part of the current NRSA design.

A.1.5 *Multi-density categories*

A complex unequal probability selection process was used in each of the two components of the survey design. They are described separately.

1.1.1.1 NRSA design

Unequal probability categories are defined separately for wadeable streams (1^{st} to 4^{th} order) and non-wadeable rivers (5^{th} to 10^{th} order). Note that "wadeable" and "non-wadeable" are used to designate Strahler order classes and do not imply that the streams will actually be wadeable or non-wadeable. The expected sample size is 450 for wadeable streams and 900 for non-wadeable rivers.

For the wadeable stream category, within each state unequal selection probabilities were defined for 1^{st}, 2^{nd}, 3^{rd}, and 4^{th} order streams so that an equal number of sites would occur for each order. Then these unequal selection probabilities were adjusted by WSA nine aggregated ecoregion categories so that an equal number of sites would occur in each WSA nine aggregated ecoregion category.

For the non-wadeable river category, unequal selection probabilities were defined for 5^{th}, 6^{th}, 7^{th}, and 8^{th}+ order rivers so that the expected number of sites would be 350, 275, 175, and 100 sites, respectively. Then these unequal selection probabilities were adjusted by WSA nine aggregated ecoregion categories so that an equal number of sites would occur in each WSA nine aggregated ecoregion category.

Given these initial selection probabilities, the expected number of urban and non-urban sites was calculated to determine if at least 150 urban sites would be selected. Over 150 urban sites were expected, so no additional adjustment was required to satisfy the urban design requirement.

The final adjustment of the selection probabilities was to adjust them to minimize the range in the number of sites across the 48 states while still meeting the other design requirements. Given a total of 1,350 sites for the NRSA design, each state would sample 28 sites. This could not be achieved, although the range was able to be decreased.

1.1.1.2 WSA_Revisit design

The Wadeable Streams Assessment sampled 1,390 sites between 2000 and 2004. To estimate change, 450 of these sites wereevisited as part of the 2008–2009 Rivers and Streams assessment. The revisit design selects the 450 sites using unequal selection probabilities. Initially, all sites were assigned an equal selection probability of 1.

First, four intensification study regions were sampled as part of the WSA. These regions are the Wenatchee Watershed in Washington, Lower John Day and Deschutes watersheds in Oregon, Northern California coastal watersheds, and southern California coastal give the expected number of sites within a study region if a state-wide survey design was done without intensification.

Second, the density of sites sampled for the EMAP-West portion of the WSA was greater than for the 36 eastern states. The selection probabilities were reduced for EMAP-West states to adjust for this. The density of sites in the Southern Appalachian aggregated ecoregion was less than in other eastern aggregated ecoregions as a result of the site replacement process used in the WSA. The selection probabilities were increased for these sites as well. The latter also ensured that the final weights for these sites were not extreme.

Third, the selection probabilities developed above were adjusted to achieve approximately an equal number of sites across all nine WSA aggregated ecoregions.

Fourth, the overall weight, inverse of selection probability, was calculated by multiplying the original WSA weight by the inverse of the above selection probability. This accounts for the fact that the WSA_Revisit design is a two-stage sample of wadeable streams.

WSA_Revisit design weights and NRSA design weights associated with wadeable streams will have to be adjusted to account for the fact that they are two independent survey designs of wadeable streams for the 48 states. This will be done after the sites are evaluated and sampled.

1.1.1.3 State designs

For any state that has a current, compatible state-wide probability design that covers all flowing waters, an option is provided to use its sites instead of the flowing water design sites. For the option to be exercised for a state, (1) the state's design must be a probability survey design, (2) its target population of streams and rivers must include the target population for the NRSA, (3) its sample frame must include the NRSA sample frame, and (4) its design must be implemented statewide in 2008–2009. The state must also agree to measure all the indicators included in the National Rivers and Streams Assessment using the national field and laboratory protocols.

1.1.2 ***Oversample***

No over sample sites were selected for the WSA_Revisit design. The expectation is that all, or almost all, of the 450 sites selected will be sampled given that they were sampled previously.

For the NRSA design, the oversample is nine times the expected sample size within each state. The large oversample size was used to accommodate states that may want to increase the number of sites sampled within their state for a state-level design.

1.1.3 *Site use*

Each stream/river selected to be sampled is given a unique site identification (siteID) with two parts: (1) NFW08 that identifies the sites as part of the 2008-9 National Rivers and Streams Assessment and (2) the two-letter state FIPS code followed by a number between 001 and 999 within each state. It critical that this siteID be used in its entirety to make sure that the stream and river sites are correctly identified.

Sites are organized to be used within each state. If evaluation determines that a stream or river site cannot be sampled, it is replaced by another site within the state. Sites that are coded as 1st, 2nd, 3rd, and 4th are to be replaced by over sample sites that are coded 1st, 2nd, 3rd, or 4th, ignoring order within this range. For example, a 2nd order would be replaced by either a 1st, 2nd, 3rd, or 4th order stream. Sites that are coded as 5th, 6th, 7th, 8th, 9th, or 10th order are to be replaced by oversample sites that are coded 5th, 6th, 7th, 8th, 9th, or 10th order, ignoring order within this range. For example, a 5th order river would be replaced by a 5th, 6th, 7th, 8th, 9th, or 10th order river. In each case the next lowest siteID that is within the Strahler order set is used for the replacement.

1.2 Evaluation process

The survey design weights that are given in the design file assume that the survey design is implemented as designed. Typically, users prefer to replace sites that cannot be sampled with other sites to achieve the sample size planned. The site replacement process is described above. When sites are replaced, the survey design weights are no longer correct and must be adjusted. The weight adjustment requires knowing what happened to each site in the base design and the over sample sites. EvalStatus is initially set to "NotEval" to indicate that the site has yet to be evaluated for sampling. When a site is evaluated for sampling, then the EvalStatus for the site must be changed. Recommended codes are:

EvalStatus Code	Name	Meaning
TS	Target Sampled	Site is a member of the target population and was sampled
LD	Landowner Denial	Landowner denied access to the site
PB	Physical Barrier	Physical barrier prevented access to the site
NT	Non-Target	Site is not a member of the target population
NN	Not Needed	Site is a member of the over sample and was not evaluated for sampling
Other codes		Other codes are often useful. For example, rather than use NT, may use specific codes indicating why the site was non-target.

1.3 Statistical analysis

Any statistical analysis of data must incorporate information about the monitoring survey design. In particular, when estimates of characteristics for the entire target population are computed, the statistical analysis must account for any stratification or unequal probability selection in the design. Procedures for doing this are available from the Aquatic Resource Monitoring Web page given in the bibliography. A statistical analysis library of functions is available from the Web page to do common population estimates in the statistical software environment R.

1.4 Literature cited

Diaz-Ramos, S., D. L. Stevens, Jr, and A. R. Olsen. 1996. EMAP Statistical Methods Manual. EPA/620/R-96/002, U.S. Environmental Protection Agency, Office of Research and Development, NHEERL-Western Ecology Division, Corvallis, Oregon.

Horn, C.R. and Grayman, W.M. (1993) Water-quality modeling with EPA reach file system. Journal of Water Resources Planning and Management, 119, 262-74.

Stevens, D.L., Jr. 1997. Variable density grid-based sampling designs for continuous spatial populations. Environmetrics, 8:167-95.

Stevens, D.L., Jr. and Olsen, A.R. 1999. Spatially restricted surveys over time for aquatic resources. Journal of Agricultural, Biological, and Environmental Statistics, 4:415-428

Stevens, D. L., Jr., and A. R. Olsen. 2003. Variance estimation for spatially balanced samples of environmental resources. Environmetrics **14**:593-610.

Stevens, D. L., Jr., and A. R. Olsen. 2004. Spatially-balanced sampling of natural resources in the presence of frame imperfections. Journal of American Statistical Association:99:262-278.

Strahler, A.N. 1957. Quantitative Analysis of Watershed Geomorphology. Trans. Am. Geophys. Un. 38,913-920.

Reference Condition for the NRSA 2008–2009

2.1 Background information

In order to assess current ecological condition, it is necessary to compare measurements today to an estimate of expected measurements in a less-disturbed situation. Because of the difficulty in establishing pristine conditions for many indicators, the NRSA used "Least-Disturbed Condition" as the reference condition. Least-Disturbed Condition can be defined as the best available chemical, physical, and biological habitat conditions given the current state of the landscape. Reference thresholds describe the sites whose condition is "the best of what's left." Data from reference sites were used to select metrics for IBIs, develop O/E models, and define the ecoregion-specific condition class thresholds used in the NRSA.

2.2 Methodology

2.2.1 *Sources of reference sites*

The fish and macroinvertebrate reference sites used in the NRSA came from four major activities:

1. Sites sampled during the NRSA using consistent sampling protocols and analytical methods that were screened to meet ecoregion-specific physical and chemical criteria. These included both sites selected randomly from the probability sample and sites hand-picked by best professional judgment and sampled using NRSA methods as part of the NRSA. Sites sampled as hand-picked, targeted reference sites for the NRSA were identified as reference via a three tiered approach. First, sites throughout the country that were submitted as least-disturbed by states, academics, USGS, and EPA Regions were screened using a quantitative disturbance score for the local watershed (the area draining to the reach segment). Sites were then sent to the EPA Landscape Ecology Lab for a quantitative disturbance score for the cumulative watershed (includes the reach and all upstream reaches). Finally, the top 300 sites that ranked using a visual assessment of disturbance at the 1:24,000 and 1:3,000 scales. 200 sites were selected that covered the nine ecoregions and two resource types and ranked high across all screens.

2. In addition to the sites sampled in the NRSA, as part of the NRSA data analysis process, we obtained possible reference site external data from NAWQA, EPA Region 7, the State of Wisconsin, and the State of Oklahoma. These data included fish and macroinvertebrate assemblage data as well as physical and chemical habitat data.

3. Benthic reference site data from 1,655 wadeable stream sites was available from the 2006 EPA Wadeable Streams Assessment (WSA). In the WSA, reference sites were obtained from two different approaches: first by screening the WSA survey data for physical and chemical criteria in the same manner described in #1 above, and second from macroinvertebrate sample data provided by other agencies, universities, or states

from sites that were deemed to be suitable as reference sites by best professional judgment. These sites either were sampled with the same methodology as the WSA or had field and lab protocols with enough similarities that the data analysis group determined that the data were comparable. The reference sites from this second approach were only used in developing a MMI for benthic samples, not for setting nutrient thresholds. The WSA reference site screening process and data sources are described in detail in Herlihy et al. (2008). The first two data columns in Table B-1 summarize the number of available WSA macroinvertebrate reference sites by ecoregion.

4. Fish reference site data from stream and river sites used by Herlihy et al. (2006) in a national analysis of fish assemblage data. The screening process used to define reference sites is described in Herlihy et al. (2006) and defined in detail in Appendix 1 of that document. The Herlihy et al. (2006) study only used the first two years of data from EMAP-West. The last three years of the data from EMAP-West was also available so that reference fish data was used as well. Final numbers of reference sites and screening used to refine the fish reference population are outlined in that section of the technical support documents.

Table B-1. Macroinvertebrate reference sites available for use in the NRSA.

Ecoregion	WSA Activities		NRSA Activities		Total
	WSA—External	WSA—Screened	NRSA—External	NRSA—Screened	
Northern Appalachians (NAP)	114	27	2	37	180
Southern Appalachians (SAP)	370	35	22	38	465
Coastal Plain (CPL)	112	15	3	46	176
Upper Midwest (UMW)	68	12	38	30	148
Temperate Plains (TPL)	124	38	50	22	234
Northern Plains (NPL)	10	18	3	47	78
Southern Plains (SPL)	56	21	51	34	162
Western Mountains (WMT)	335	129	4	40	508
Xeric Region (XER)	132	39	2	33	206
Total	1,321	334	175	327	2,157

2.2.2 *Screening NRSA data for reference condition*

To identify reference sites by screening the NRSA data, we used the chemical and physical data collected at each site (e.g., nutrients, turbidity, acidity, riparian condition) to determine whether any given site is in least-disturbed condition for its ecoregion. In the NRSA, eight physical and chemical parameters were used to screen for reference sites, total N, total P, chloride, sulfate, acid neutralizing capacity, turbidity, % fine substrate, and riparian disturbance index. If a site exceeded the screening value for any one stressor it was dropped from reference consideration. Given that expectations of least-disturbed condition vary across ecoregions, the criteria values for exclusion varied by ecoregion. The nine aggregate level III ecoregions developed for the WSA assessment were used to regionalize reference conditions. Ecoregional specific screening criteria are listed in Table B-2. The Western Mountains ecoregion was broken into three finer-scale ecoregion subgroups for screening to match what was done in EMAP-West which was done at a somewhat finer spatial scale.

In addition to the sites sampled in the NRSA, we obtained possible reference site external data from four other agencies. Data from these external surveys were screened for physical and chemical criteria using the same criteria used for NRSA sample sites in Table B-2 using whatever screening data were available in each survey.

All sites in the NRSA (both probability and hand-picked, boatable and wadeable) and the added external data that passed all criteria were considered to be candidate reference sites for the NRSA assessment. The number of reference sites that passed this screening is summarized in Table B-1. These reference sites include both fish and macroinvertebrate data. Note that the NRSA did not use data on the biological assemblages themselves for any screening as these are the primary components of the stream and river ecosystems being evaluated, and to use them would constitute circular reasoning.

Table B-2. Criteria for eight chemical and physical habitat filters used to identify the candidate least-disturbed reference sites in NRSA for each of the nine aggregate ecoregions.

Note that RBP physical habitat score was used as a filter in WSA but was not available in the NRSA data to use as a screen. The six ecoregions in the top half of the table were used in WSA and reported in Herlihy et al. (2008), the ecoregions in the bottom half of the table were screened using criteria developed in EMAP-West. ANC = acid neutralizing capacity, DOC = dissolved organic C.

Filter criterion	NAP	SAP	CPL	UMW	TPL	SPL
Total P (µg/L)	>20	>20	>75	>50	>100	>150
Total N (µg/L)	>750	>750	>2500	>1000	>3000	>4500
Cl^- (µeq/L)	>250[a]	>200	–	>300	>2000	>1000
SO_4^{2-} (µeq/L)	>250	>400	>600	>400	–	–
ANC (µeq/L) + DOC (mg/L)[b]	<50 + <5	<50 + <5	<50 + <5	<50 + <5	<50 + <5	<50 + <5
Turbidity (NTU)	>5	>5	>10	>5	>50	>50
Riparian Disturbance Index[c]	>2	>2	>2	>2	>2	>2
% fine substrate	>25	>25	>50	>40	>80	>90

Values in red indicate a change from that used in WSA as reported in Herlihy et al., (2008).

Filter criterion	NPL	XER	WMT-SW[e]	WMT-SRock[e]	WMT-Nrock/Pacific[e]
Total P (µg/L)	>150	>50	>50	>25	>25
Total N (µg/L)	>4500	>1500	>750	>750	>750
Cl^- (µeq/L)	>1000	>1000	>300	>200	>200[a]
SO_4^{2-} (µeq/L)	–	–	–	>200	>200
ANC (µeq/L) + DOC (mg/L)[b]	<50 + <5	<50+<5	<50 + <5	<50 + <5	<50 + <5
Turbidity (NTU)	>50	>25	>5	>5	>5
Riparian Disturbance Index	>2	>1.5	>0.5/>1.5[d]	>1/>1.5[d]	>0.5/>1.5[d]
% fine substrate	>90	>50	>15	>15	>15

– indicates filter criterion was not used in that ecoregion.

[a] Cl^- criterion not applied in Northeastern Coastal Zone (ecoregion 59) or Coast Range (ecoregion 1) sites

[b] Filter was specific for inorganic acidity; site had to exceed both criteria to fail

[c] Riparian disturbance index variable name is W1_HALL in physical habitat database.

[d] Wadeable stream/Boatable river criteria. Different criteria were used by stream size in the Western Mountains.

[e] To match screening criteria to what was done in the EMAP-West component of WSA, the Western Mountains ecoregion was divided into three subgroups: SW = Southwestern Mountains (Omernik level III codes 8 and 23, Southern California Mts., and Arizona/New Mexico Mts.), SRock = Southern Rockies (Omernik 19 and 21, Southern Rockies and Wasatch/Uintas), and NRock/Pacific = Northern Rockies and Pacific Mountains (all other WMT level III ecoregions).

2.2.3 ***Final combined reference site screen***

As a final screen, all of the NRSA screened reference sites, and those provided by WSA and any other source, were screened for the influence of dams and adjacent land use. Three additional landscape-GIS screening criteria were applied to the selected physiochemical screened reference sites. These screens included dam influence index, urbanization influence, and agricultural influence.

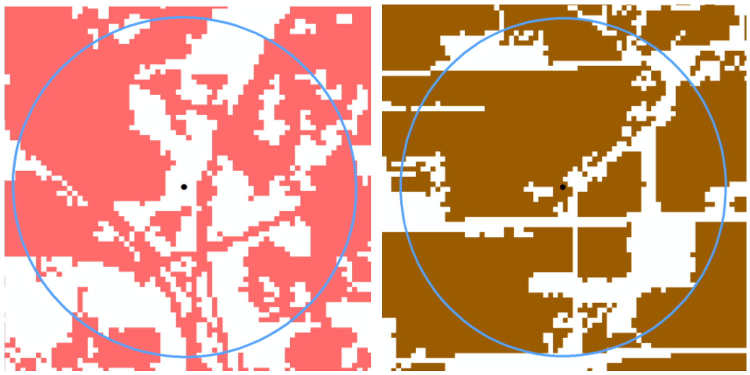
The dam influence index (DII) was used to assess the influence of upstream dams and the largest reservoir on NRSA reference sites. Any watershed boundaries that had a maximum distance of less than 200 km upstream of the sampling point were completely assessed, any watershed with a distance greater than 200 km upstream of the sample point, had a wedge shaped area assessed until 200 km upstream was reached. For all watersheds and wedges assessed, a calculation of the volume of the largest reservoir, the number of dams, and an index that weighted the maximum reservoir volume within the watershed or wedge by its proximity to the sample point was conducted. Each upstream reservoir was inversely weighted by its upstream flow distance from the sample point as:

,

where D_{flow} is the flow distance to the sample site, and $D_{efolding}$ is an e-folding value that determines the rate at which the weight exponentially decreases (here 100 km). DII equals the largest distance-weighted volume within the watershed:

where D_i = reservoir volume (km^3). The threshold for dropping a potential reference site was a DII value equal to or greater than one.

Percent urbanization and agricultural influence was assessed within a 1 km^2 area around the mid-point of the sampled stream segment. To conduct this analysis a 1 km^2 radius buffer around the mid-point was overlaid onto the National Land Cover Dataset (NLCD) to calculate the percentage of urban land cover and percent row crop, as defined by the NLCD (Figure B-1). The threshold for dropping a potential reference sites was any greater than 5% urban land cover and 15% agricultural (row crop) land cover.



A B

Figure B-1. Examples of percent urban (A, 60%) and row crop (B, 72%) from NLCD.

2.3 Literature cited

Herlihy, A.T., R.M. Hughes, and J.C. Sifneos. 2006. National clusters of fish species assemblages in the conterminous United States and their relationship to existing landscape classification schemes. pp. 87-112. In R.M. Hughes, L. Wang, and P.W. Seelbach (eds.), Influences of Landscapes on Stream Habitats and Biological Assemblages. American Fisheries Society Symposium 48, Bethesda, Maryland

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

Benthic Macroinvertebrate Assemblage for NRSA 2008–2009

3.1 Background information

The taxonomic composition and relative abundance of different taxa that make up the benthic macroinvertebrate assemblage present in a stream have been used extensively in North America, Europe, and Australia to assess how human activities affect ecological condition (Barbour et al. 1995, 1999; Karr and Chu 1999). Two principal types of ecological assessment tools to assess condition based on benthic macroinvertebrates are currently prevalent: multimetric indices and predictive models of taxa richness. The purpose of these indicators is to present the complex community taxonomic data represented within an assemblage in a way that is understandable and informative to resource managers and the public. The following sections provide a general overview of the approaches used to develop ecological indicators based on benthic macroinvertebrate assemblages, followed by details regarding data preparation and the process used for each approach to arrive at a final indicator.

3.1.1 *Overview of the IBI and O/E predictive model approaches*

Multimetric indicators have been used in the U.S. to assess condition based on fish and macroinvertebrate assemblage data (e.g., Karr and Chu, 1999; Barbour et al., 1999; Barbour et al., 1995). The multimetric approach involves summarizing various assemblage attributes (e.g., composition, tolerance to disturbance, trophic and habitat preferences) as individual "metrics" or measures of the biological community. Candidate metrics are then evaluated for various aspects of performance and a subset of the best performing metrics are then combined into an index, referred to as a multimetric index or MMI. For NRSA, the MMI developed in the WSA was used to generate the population estimates used in the assessment. The WSA MMI is detailed in Stoddard et al. (2008).

The predictive model approach was initially developed in Europe and Australia, and is becoming more prevalent within the U.S. The approach estimates the expected taxonomic composition of an assemblage in the absence of human stressors (Hawkins et al., 2000; Wright, 2000), using a set of "least-disturbed" sites and other variables related natural gradients (such as elevation, stream size, stream gradient, latitude, longitude). The resulting models are then used to estimate the expected taxa composition (expressed as taxa richness) at each stream site sampled. The number of expected taxa actually observed at a site is compared to the total number of expected taxa as an observed:expected ratio (O/E index). Departures from a ratio of 1.0 indicate that the taxonomic composition in a stream sample differs from that expected under less disturbed conditions.

3.2 Data preparation

3.2.1 *Standardizing counts*

The number of individuals in a sample was standardized to a constant number to provide an adequate number of individuals that was the same for the most samples and that could be used for both multimetric index development and O/E predictive modeling index. A subsampling technique involving random sampling without replacement was used to extract a true “fixed count” of 300 individuals from the total number of individuals enumerated for a sample (target lab count was 500 individuals). Samples that did not contain at least 300 individuals were used in the assessment because low counts can indicate a response to one or more stressors. Only those sites with at least 250 individuals, however, were used as reference sites.

3.2.2 *Operational taxonomic units*

For the predictive model approach, it was necessary to combine taxa to a coarser level of common taxonomy. This new combination of taxa is termed an “operational taxonomic unit” or OTU, and results in fewer taxa than are present in the initial benthic macroinvertebrate count data.

3.2.3 *Autecological characteristics*

Autecological characteristics refer to specific ecological requirements or preferences of a taxon for habitat preference, feeding behavior, and tolerance to human disturbance. These characteristics are prerequisites for identifying and calculating many metrics. A number of state/regional organizations and research centers have developed autecological characteristics for benthic macroinvertebrates in their region. For the WSA and NRSA, a consistent “national” list of characteristics that consolidated and reconciled any discrepancies among the regional lists was needed before certain biological metrics could be developed and calibrated and an MMI could be constructed.

Members of the data analysis group pulled together autecological information from five existing sources: the EPA Rapid Bioassessment Protocols document, the National Ambient Water Quality Assessment (NAWQA) national and northwest lists, the Utah State University list, and the EMAP Mid-Atlantic Highlands (MAHA) and Mid-Atlantic Integrated Assessment (MAIA) list. These five were chosen because they were thought to be the most independent of each other and the most inclusive. A single national-level list was developed based on the following decision rules:

3.2.4 *Tolerance values*

Tolerance value assignments followed the convention for macroinvertebrates, ranging between 0 (least tolerant or most sensitive) and 10 (most tolerant). For each taxon, tolerance values from all five sources were reviewed and a final assignment made according to the following rules:

1. If values from different lists were all <3 (sensitive), final value = mean.
2. If values from different lists were all >3 and <7 (facultative), final value = mean.
3. If values from different lists were all >7 (tolerant), final value = mean.
4. If values from different lists spanned sensitive, facultative, and tolerant categories, best professional judgment was used, along with alternative sources of information (if available) to assign a final tolerance value.
5. Tolerance values of 0 to ≤3 were considered "sensitive" or "intolerant." Tolerance values ≥7 to 10 were considered "tolerant," and values in between were considered "facultative."

3.2.5 *Functional feeding group and habitat preferences*

In many cases, there was agreement among the five data sources. When discrepancies in functional feeding group (FFG) or habitat preference ("habit") assignments among the five primary data sources were identified, a final assignment was made based on the most prevalent assignment. In cases where there was no prevalent assignment, the workgroup examined why disagreements existed, flagged the taxon, and used best professional judgment to make the final assignment.

3.3 Multimetric Index development

3.3.1 *Regional multimetric development*

The autecology and taxonomic resolution used in WSA was applied to the NRSA macroinvertebrate 300 fixed count data to calculate the community metrics used to calculate the MMI. In the WSA, a best ecoregional MMI was developed by summing the six metrics that performed best in that ecoregion (the national aggregate nine ecoregions). Each metric was scored on a 0–10 scale and then summed and normalized to a 0–100 scale to calculate the final MMI. Details of this process are described in Stoddard et al. (2008). The final metrics used in each ecoregion are summarized in Table C-1. The NRSA MMI was calculated in the same way as the WSA MMI. Based on NRSA revisit data, the MMI had a S:N ratio of 2.8 and a pooled standard deviation of 10.0 (out of 0–100).

Table C-1. Six benthic community metrics used in for the NRSA and WSA MMI in each of the nine aggregate ecoregions.

Metric	NAP	SAP	CPL	UMW	TPL	NPL	SPL	WMT	XER
EPT_Percent Distinct Taxa	X					X		X	
EPT_Percent Individuals					X		X		
Non-Insect % Distinct Taxa									X
Non-Insect % Distinct Individuals			X						
Ephemeroptera % Distinct Taxa		X							
Chironomid % Distinct Taxa				X					
Shannon Diversity		X	X	X	X		X		
% Individuals in top 5 taxa	X							X	X
% Individuals in top 3 taxa						X			
Scraper Richness	X	X			X	X	X	X	X
Shredder Richness			X	X					
Burrower % Distinct Taxa		X		X		X	X		
Clinger % Distinct Taxa	X		X					X	X
Clinger Distinct Taxa Richness					X				
Ephemeroptera Distinct Taxa Richness					X				
EPT Distinct Taxa Richness	X	X	X	X			X	X	X
Total Distinct Taxa Richness						X			
Intolerant Richness						X	X		
Tolerant % Distinct Individuals		X	X					X	X
PTV 0-5.9% Distinct Taxa	X								
PTV 8-10% Distinct Taxa				X	X				

3.3.2 *Modeling of MMI condition class thresholds for the Wadeable Streams Assessment*

Previous large-scale assessments have converted MMI scores into classes of assemblage condition by comparing those scores to the distribution of scores observed at least-disturbed reference sites. If a site's MMI score was less than the 5th percentile of the reference distribution, it was classified as in "poor" condition; scores between the 5th and 25th percentile were classified as "fair," and scores in the 25th percentile or higher were classified as "good." This approach assumes that the distribution of MMI scores at reference sites reflects an approximately equal, minimum level of human disturbance across those sites. But this assumption did not appear to be valid for some of the nine WSA regions, which was confirmed by state and regional parties at meetings to review the draft results.

For the WSA, the project team performed a principal components analysis (PCA) of NINE habitat and water chemistry variables that had originally been used to select IBI reference sites. The first principal component (Factor 1) of this PCA represented a generalized gradient of human disturbance. MMI scores at the reference sites were weakly, but significantly, related to this disturbance gradient in five of the nine aggregate regions. Thus, MMI reference distributions from these regions are biased downward, because they include somewhat disturbed sites which have lower MMI scores. As part of the WSA, Herlihy et al. (2008) developed a process that used the PCA disturbance gradient scores to reduce the effects of disturbance within the reference site population. The process used multiple regression modeling to develop adjusted thresholds analogous to the 5th and 25th percentiles of reference sites in each ecoregion. These adjusted thresholds were used in the WSA and the same thresholds were used in the NRSA to define good, fair, and poor condition based on the benthic MMI. The process for calculating these adjusted thresholds is detailed in Herlihy et al. (2008), and the threshold values for each ecoregion used in WSA and the NRSA report are given in Table C- 2.

Table C-2. Threshold values for the nine regional benthic MMIs. Any site with an MMI score that was not good or poor was considered "fair."

Ecoregion	Good Threshold	Poor Threshold
CPL	≥56	<42
NAP	≥63	<49
NPL	≥55	<41
SAP	≥51	<37
SPL	≥44	<30
TPL	≥45	<31
UMW	≥48	<34
WMT	≥54	<40
XER	≥53	<40

3.4 O/E: predictive (RIVPACS) models

In addition to the benthic macroinvertebrate MMI approach, predictive O/E modeling was used to assess benthic macroinvertebrate condition for the NRSA. The O/E model compares the observed benthic assemblage at a site to an expected assemblage derived from a population of reference sites. Stressors and anthropogenic impacts lead to a reduction in the number of taxa that are expected to be present under reference conditions. The predictive model approach is used by several states and is a primary assessment tool of Great Britain and Australia.

The O/E ratio predicted by the model for any site expresses the number of taxa found at that site (O), as a proportion of the number that would be expected (E) if the site was in least-disturbed condition. Ideally, a site in reference condition has O/E = 1.0. An O/E value of 0.70 indicates that 70% of the "expected" taxa at a site were actually observed at the site. This is interpreted as a 30% loss of taxa relative to the site's predicted reference condition. However, O/E values vary among reference sites themselves, around the idealized value of 1.0, because such sites rarely conform to an idealized reference condition, and because of model error and sampling variation. The standard deviation of O/E (Table C-3) indicates the breadth of O/E variation at reference sites. Thus, the O/E value of an individual site should not be interpreted as (1 – taxa loss) without taking account of this variability in O/E. Individual O/E values are most reliably interpreted relative to the entire O/E distribution for reference sites.

A nationally distributed collection of reference sites was first identified, drawn from a pool of sites whose macroinvertebrates were sampled using EMAP protocols. This pool included only NRSA, WSA, EMAP-West, STAR-Hawkins, USGS NAWQA, and MAHA/MAIA sites. One hundred reference sites were set aside to validate the models, and the remaining reference sites were used to calibrate the models (Table C-3). Each site contributed a single sampled macroinvertebrate assemblage to model calibration and validation. Each sampled macroinvertebrate assemblage comprising more than 300 identified individuals was randomly subsampled to yield 300 individuals. 300-count subsamples were used to build models and assess all NRSA sites.

The predictive modeling approach assumes that expected assemblages vary across reference sites throughout a region, due to natural (non-anthropogenic) environmental features such as geology, soil type, elevation, and precipitation. To model these effects, the approach first classifies reference sites based on similarities of their macroinvertebrate assemblages (Table C-3). A random forest model is then built to predict the membership of any site in these classes, using natural environmental features as predictor variables (Table C-3). The predicted occurrence probability of a reference taxon at a site is then predicted to be the weighted average of that taxon's occurrence frequencies in all reference site classes, using the site's predicted group membership probabilities in the classes as weights. Finally, E for any site is the sum, over a subset of reference taxa, of predicted taxon occurrence probabilities. O is the number of taxa in that subset that were observed to be present at the site. The subset of reference taxa used for any site was defined as those taxa with predicted occurrence probabilities exceeding 0.5 at that site.

Final predictive models performed better than corresponding null models (no adjustment for natural-factor effects), as judged by their smaller standard deviation of O/E across calibration sites (Table C-3).

Similar to the IBI, two scaled approaches were used to develop the O/E model. A national model was initially developed to predict taxa loss at sites. Three models were developed for NRSA usage, together covering the contiguous USA (Table C-3). The regional models performed better, and were used in the NRSA to predict taxa loss at the sites.

Table C-3. NRSA predictive models.

Model Name	Eastern Highlands	Plains and Lowlands	West
Regions covered	NAP, SAP	CPL, UMW, TPL, NPL, SPL	WMT, XER
Number of calibration sites	297	241	659
Number of validation sites	31	21	48
Number of site classes	17	16	
Random Forest predictor variables	Predicted mean summer stream temperature, watershed area, watershed mean minimum annual temperature, predicted mean annual stream temperature, watershed mean annual temperature, watershed mean minimum precipitation	Predicted mean annual stream temperature, watershed mean date of last freeze, watershed mean soil permeability, watershed mean runoff, watershed maximum elevation	Watershed area, watershed mean annual temperature, watershed mean precipitation accumulation, predicted mean annual stream temperature, watershed mean maximum temperature, watershed mean elevation
Standard deviation of O/E at calibration sites: -- Predictive model -- Null model	 0.18 0.22	 0.23 0.26	 0.18 0.25

3.5 Literature cited

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

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

L.L. Yuan, C.P. Hawkins, J. Van Sickle. 2008. Effects of regionalization decisions on an O/E index for the US national assessment. Journal of the North American Benthological Society 27:892-905.

Fish Community Assemblage for NRSA 2008–2009

4.1 Background information

Fish assemblages in streams and rivers offer several unique advantages to assess ecological condition, based on their mobility, longevity, trophic relationships, and socioeconomic importance (Barbour et al. 1999, Roset et al. 2007). For fish assemblages, assessing ecological condition has generally been based the development and use of multimetric indices (MMIs), which are derivations of the original Index of Biotic Integrity (IBI) developed by Karr and others (Karr 1981, Karr et al. 1986, Karr 1991, 1999, Karr and Chu 2000). There are numerous examples of MMIs developed for fish assemblages in smaller streams (e.g., Bramblett et al. 2005) as well as for larger rivers (Lyons et al. 2001, Emery et al. 2003, Mebane et al. 2003, Pearson et al. 2011).

Recently, MMIs for fish assemblages have been developed based on applying the techniques used to develop predictive models of taxonomic composition (e.g., Hawkins 2006, Meador and Carlisle 2009) to metrics instead of taxa (Pont et al. 2007, Pont et al. 2009). This approach essentially provides an estimate of expected condition (in terms of metric values) at individual sites, rather than using a set of regional reference sites to define expected values for a particular metric, Hawkins et al. (2010a) concluded that the combined approach resulted in MMIs that performed better in terms of their ability to discern deviation from expected condition.

For NRSA, we developed multimetric indices for fish assemblages (FMMIs) using the combined approach (modeling expected values of metrics). We developed separate FMMIs for each of the three climatic regions (Eastern Highlands, Plains and Lowlands, and West; Figure D-1).

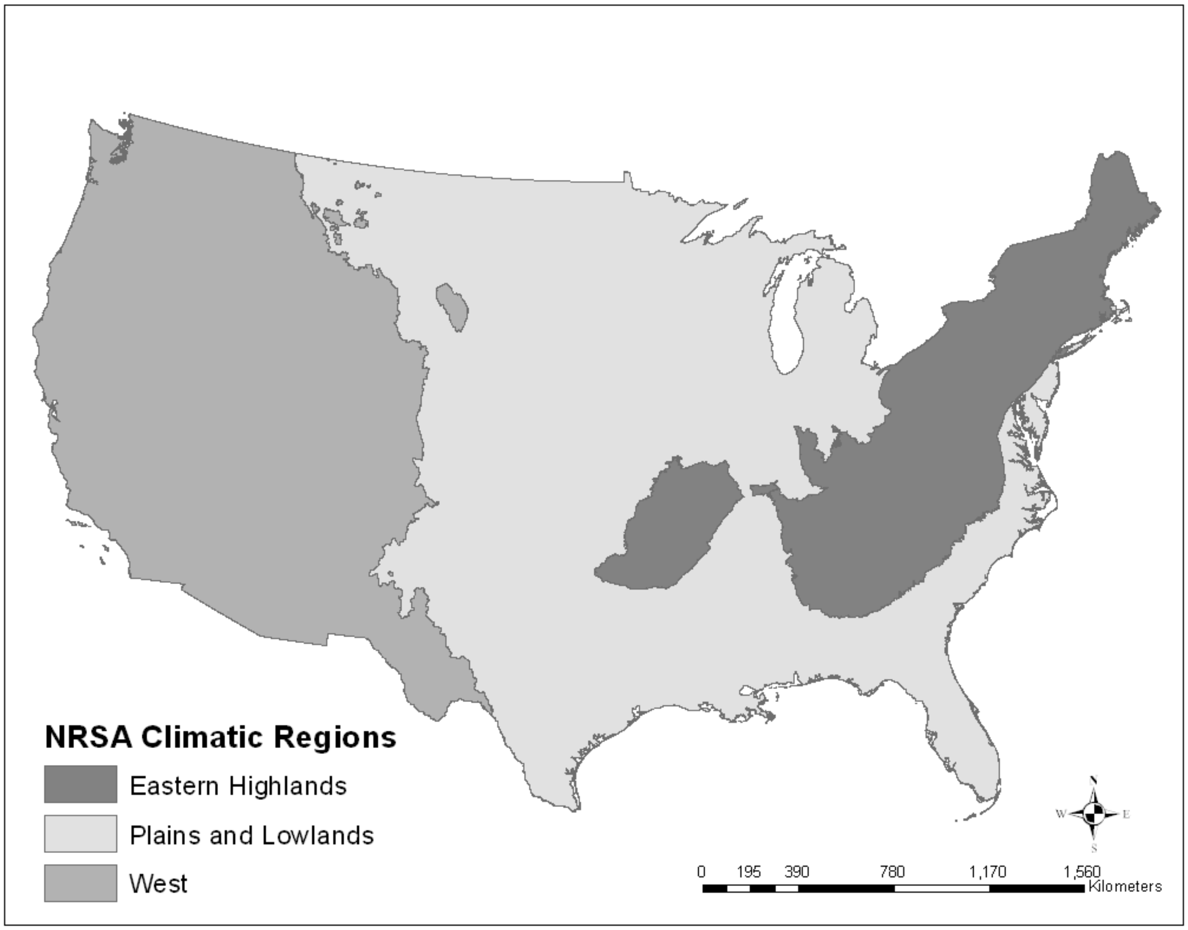


Figure D-1. Aggregated Omernik ecoregions used for NRSA fish MMI development. Separate MMIs were developed for each of the three climatic regions.

4.2 Methods

4.2.1 *Field methods*

Collection methods for fish are described in the NRSA field operations manual (EPA 2009). Three variants of the basic sampling protocol (using electrofishing) were used depending on the width of the stream and whether or not it was wadeable. For wadeable streams less than 12.5 meters wide, 40 channel widths were sampled for fish. For larger wadeable streams (>12.5 meters wide), 500 meters or 20 channel widths were sampled (whichever was longer). For non-wadeable streams and rivers, at least 20 channel widths were sampled. At large wadeable and non-wadeable sites, sampling continued past the established reach length until 500 individuals were collected.

4.2.2 *Counting, taxonomy, and autecology*

Fish were tallied and identified in the field, then released alive unless used for fish tissue or vouchers. Voucher specimens were collected if field identification could not be accomplished.

Voucher samples were also collected at 10% of sites for each taxonomist. All names submitted on field data forms were reviewed and revised when necessary to create a listing of nationally consistent common and scientific names. Where possible, taxonomic names (common and scientific) were based on Nelson et al. (2004). The online database FishBase (http://www.fishbase.org) served as a secondary source of taxonomic names. In rare cases, a journal article of a newly described species was used. Collection maps for each taxon were prepared and compared to published maps in Page and Burr (1991). A total of 631 unique taxa were identified, excluding unknowns, hybrids, and amphibians.

Each taxon was characterized for several different autecological traits, based on available sources of published information. Traits included habitat guilds (lotic habitat and temperature), trophic guild, reproductive guild, migration pattern, and tolerance to human disturbance. A list of all fish taxa and their associated autecological assignments are available as a tab-delimited data file that will be available on the NRSA website in December 2012.

Assignments of native status were made at the scale of 8-digit Hydrologic Unit Codes (HUC). Published sources (USGS Nonindigenous Species database and NatureServe) were used as the basis for assigning a taxon collected at a particular site as being native or introduced.

Because fish collected at a site cannot always be confidently identified to species, there is a risk of inflating the number of species actually collected. For each sample, we reviewed the list of taxa to determine whether they were represented at more than one level of resolution. For example, if an “unknown catostomus” was collected, and it was the only representative of the genus at the site, we assigned it as distinct. If any other species of the genus were collected, then we considered the unknown as not distinct. We used only the number of distinct taxa in the sample to calculate any metrics based on species richness.

4.3 Fish Multimetric Index development

We used a consistent process to develop an FMMI for each of the three climatic regions. For each candidate metric, we applied the set of predictor variables to a set of reference sites using a random forest model (Cutler et al. 2007, Hawkins et al. 2010a). The model provided expected values for the metric (i.e., under least-disturbed conditions) given the particular values of the predictor variables. This approach served to help remove the effects of natural gradients on metric response, which are often confounded with disturbance gradients when expected values for a metric are based solely on a set of regional reference sites (Hawkins et al. 2010a). The model for each metric was then applied to the entire set of sites, and the residuals (deviation from predicted) were used as the response value for the metric. We evaluated each metric for its responsiveness to disturbance, i.e., the ability to discern between least-disturbed (reference sites) and more highly disturbed sites (following Stoddard et al. 2008). We then selected metrics representing different dimensions of assemblage structure or function to include in the FMMI based on responsiveness and lack of correlation with other metrics, again following Whittier et al. (2007) and Stoddard et al. (2008).

4.3.1 *Reference sites for fish*

We modified the base list of reference sites determined for NRSA to eliminate additional fish samples that might not be representative of least-disturbed conditions (Table D-1). The final set of 329 reference sites used for the FMMI development are listed in Appendix D-1.

To validate the random forest model for each metric, we identified a random subset of reference sites (validation sites) within each climatic region and excluded them from model development. We set aside 30 validation sites in the Eastern Highlands, 36 sites in the Plains and Lowlands, and 19 sites in the West region. Applying the model to these reference sites should produce MMI scores that are similar to those sites that were used to develop the model.

Table D-1. Criteria used to select reference sites for use in developing the FMMI.

Criteria		
Start with the base set of NRSA reference sites		
Keep sites with fish samples		
Drop sites where seining was only method of sampling		
Drop sites with insufficient sampling		
• Wadeable: Less than 50% of reach and < 500 individuals collected		
• Large Wadeable: < 500 m and < 500 individuals collected		
• Boatable: < 20 channel widths sampled		
Drop sites with sufficient sampling where < 30 individuals were collected		
Drop sites with sufficient sampling where nonnative individuals comprised >50% of total number of individuals collected		
Final Number of Reference Sites		
Eastern Highlands	154	
Plains and lowlands	180	
West	95	
Total	329	

4.3.2 *Candidate metrics*

We calculated 162 candidate metrics representing the following dimensions of fish assemblage structure and function (following Stoddard et al. 2008):

- Species richness
- Taxonomic composition
- Habitat guild
- Trophic guild
- Reproductive guild
- Migratory pattern (life history)

- Tolerance
- Nonnative species

For nearly all metrics, three variants were derived based on all taxa in the sample and for only native taxa in the sample: one based on distinct taxa richness, one based on the percent of individuals in the sample, and one based on the percent of distinct taxa in the sample. For some trophic metrics, additional variants were derived using only taxa that were not considered tolerant to disturbance. We included only those tolerance metrics based on sensitive and tolerant taxa, because the “intermediate tolerance” assignments included taxa with unknown tolerance.

4.3.3 *Predictor variables*

A total of 55 predictor variables were initially provided for all NRSA sites (including handpicked sites). These data were provided to us by Dr. Charles Hawkins and his staff at the Western Center for Monitoring and Assessment of Freshwater Ecosystems, Utah State University, Logan, Utah. These variables (Appendix D-2) represented the primary natural gradients that are believed to constrain the fish assemblage composition in the absence of human disturbance. The set of predictor variables included those related to watershed area and slope, elevation, latitude and longitude, air temperature, precipitation, and relative humidity. There were also model-derived estimates of flow, runoff, and predicted stream temperature. Many variables were estimated at the point level (at the site coordinates) and the watershed level (all values within a particular watershed were aggregated in some fashion). For the FMMI, we constrained the development of the index to only those sites that had both point and watershed level predictors, as we felt that fish might be more responsive to larger-scale natural driver variables than to more site-specific conditions.

In addition to the predictor variables provided, we calculated potential discharge (Q-POTENT_WS) as the product of runoff and watershed area, and included an aggregated ecoregion variable (FW_ECO9; see Figure D-2). We screened the set of predictor variables to eliminate those with large discrepancies in range between the set of reference sites and all other sites, and to eliminate those that had missing values for some sites, as a complete set of predictor variables is required to construct the metric models. The final set of 64 predictor variables that we used to develop the FMMI are identified in boldface type in Appendix D-3.

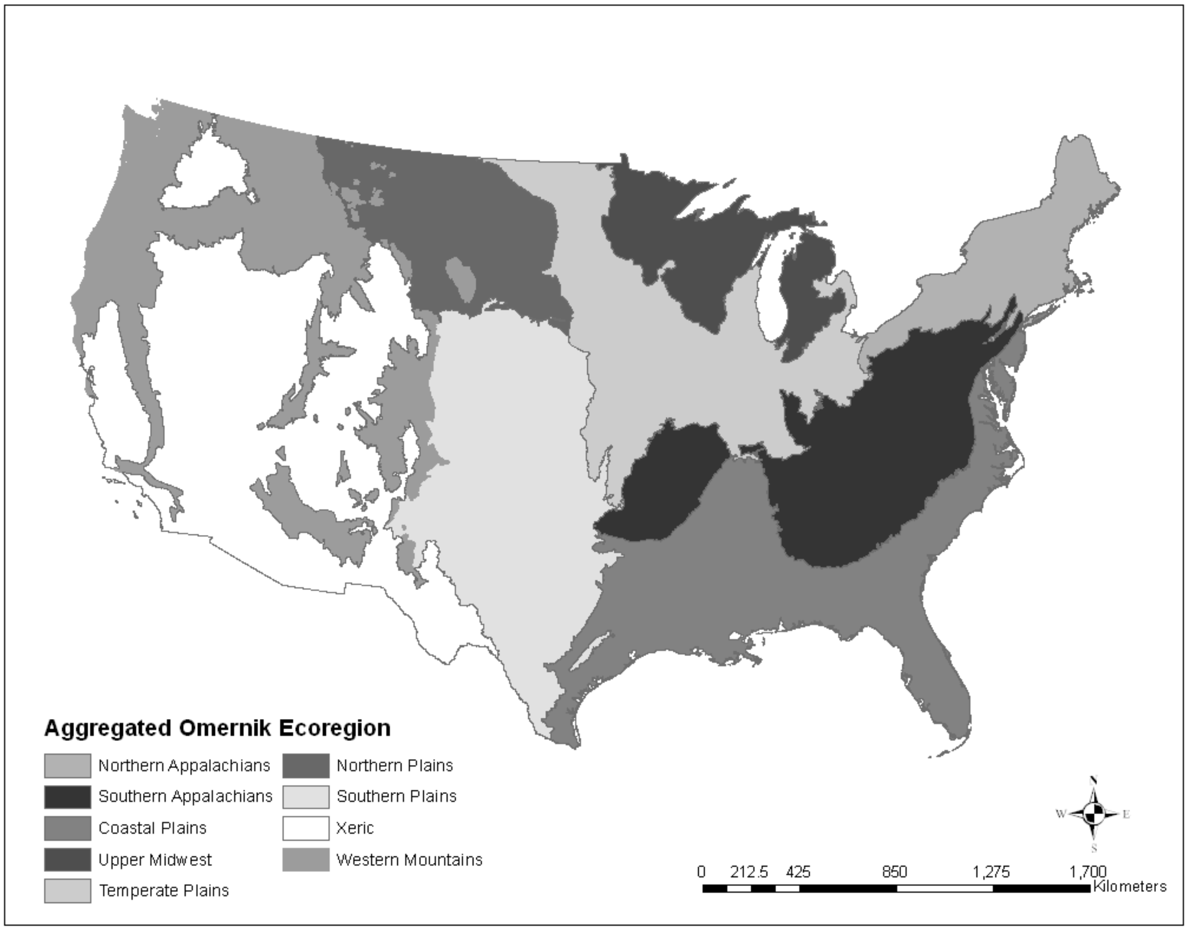


Figure D-2. Aggregated Omernik ecoregions (Level III) used as predictor variable for NRSA FMMI development and for assigning condition based on MMI score.

4.3.4 *Random forest modeling*

We used the R statistical package (version 2.13.0; R Development Core Team 2011), and the package randomForest (version 4.6-6; Liaw and Weiner 2002). We used the set of reference sites (minus those set aside for validation) to develop the metric models. The number of trees to generate was set at 500.

For metrics having a "pseudo-r^2" value >0.10, we applied the metric model to the entire set of sites, and retained the residual values as the modeled metric response value. For metrics with "pseudo-r^2" values ≤0.10, we retained the original response value.

4.3.5 *Final metric selection*

For both original and modeled metrics, the mean response values of the set of reference sites and the set of more highly disturbed sites were compared with two-sample *t*-tests (assuming unequal variances). Stoddard et al. (2008) present the advantages of using *t* values over other statistics as an indicator of metric responsiveness to disturbance. We also generated a correlation matrix (Pearson *r*) of the metrics to check for redundancy.

To select the final suite of metrics to include in the FMMI, we selected the metric with the largest value of *t* (Whittier et al. 2007, Stoddard et al. 2008). The metric with the next largest value of t that was from a different metric class than the previous metric was selected next. If this metric was uncorrelated with the first metric (r <0.7), it was retained in the final set. This process was repeated until there was one metric from each category. If no metric in a class has a *t* value >3, that category was not represented in the final suite.

Table D-2 presents the final suites of metrics selected for each of the three regional FMMIs. The Plains and Lowlands FMMI does not include a migration strategy metric, while the Eastern Highlands and West FMMIs do not include a composition metric.

We examined the variable importance plots of for each metric in the final set (produced with the R function varImpPlot) to identify the predictor variables that were the most influential in the model. We also examined partial dependence plots of the important predictor variables (produced with the R function partialPlot) to confirm the reasonableness of the relationships between predictor and metric.

Table D-2. Suite of final metrics included in each regional FMMI. Variable names are in parenthesis. Metrics in bold are modeled metrics. Values of *t* are from comparisons of mean values of reference and more highly disturbed sites.

Metric class	Eastern Highlands FMMI	Plains and Lowlands FMMI	West MMI
Richness	**Number of native sunfish species (NAT_CENTNTAX.res) *t*=-5.93**	**Number of native cyprinid species (NAT_CYPRNTAX.res) *t*=7.08**	Total number of tolerant species (TOLRNTAX) *t*=6.76
Taxonomic composition	None	% individuals that are cyprinids (CYPRPIND) *t*=4.43	None
Habitat guild	**% taxa that are native, not tolerant, and benthic (NAT_NTOLBENTPTAX.res) *t*=5.19**	**Number of native, not tolerant benthic species (NAT_NTOLBENTNTAX.res) *t*=9.35**	% taxa that are lotic (LOTPTAX) *t*=7.84
Reproductive guild	**% individuals that are lithophilic spawners (LITHPIND.res) *t*=8.24**	**% individuals that are lithophilic spawners (LITHPIND.res) *t*=8.99**	**% individuals that are lithophilic spawners (LITHPIND.res) *t*=4.03**
Trophic guild	**% taxa that are invertivores (INVPTAX.res) *t*=6.42**	**Number of native herbivore species (NAT_HERBNTAX.res) *t*=8.72**	**% taxa that are native, intolerant invertivores (NAT_INTLINVPTAX.res) *t*=5.81**
Migration strategy	**Total number of migratory taxa (MIGRNTAX.res) *t*=-4.70**	None	**% taxa that are migratory MIGRPTAX.res *t*=5.38**
Tolerance	**% taxa that are tolerant (TOLRPTAX.res) *t*=-8.38**	**% taxa that are not tolerant (NTOLPTAX.res) *t*=5.77**	% taxa that are not tolerant (NTOLPTAX) *t*=8.48
Nonnative	% taxa that are native (NAT_PTAX) *t*=3.43	**% taxa that are native (NAT_PTAX.res) *t*=4.95**	**% individuals that are native (NAT_PIND.res) *t*=10.0**

4.3.6 *Metric scoring*

Response values for each of the final suite of metrics was rescaled to a score ranging between 0 and 10. We used the 5th and 95th percentiles of all sites to set the "floor" (below which a score of 0 was assigned) and "ceiling" (above which a score of 10 was assigned) as recommended by Blocksom (2003). Response values between the floor and ceiling were assigned a score using linear interpolation.

We summed the metric scores for each site to derive the FMMI score. We then multiplied the FMMI score by (10/number of metrics) to rescale the score to range between 0 and 100 range.

4.3.7 *Sites with low fish abundance*

The target population of streams and rivers for NRSA included small headwater streams. Some very small streams may not contain fish even in the absence of human disturbance. We followed the approach described by McCormick et al. (2001) and used reference sites to estimate a drainage area below which the probability was high that no fish would be present. This approach uses the relationship between a set of four physical habitat variables that characterize habitat volume and the number of fish collected. This relationship defines a habitat volume value below which nearly all sites sampled were devoid of fish. Then this habitat volume value is related to watershed area to determine the drainage area below which streams are expected to be naturally fishless.

Figure D-3 shows the results of this analysis. The value for the habitat volume index below which almost all sites are fishless is 0.43. When habitat volume is plotted against watershed area, this value corresponds to a watershed area of approximately 2 km^2. For sites with watershed areas less than 2 km^2 where no fish were collected, we did not report the FMMI score. Otherwise, we assigned an FMMI sore of zero to sites with no fish collected.

Table D-3. Determining the minimum drainage area expected to reliably support the presence of fish. (Adapted from McCormick et al. 2001). Variable names are from the NRSA database. Scores for each metric between the upper and lower criteria were estimated by linear interpolation.

SET OF SITES
Use reference sites only (RT_NRSA_FISH="R") to minimize effects of human disturbance
HABITAT VOLUME INDEX
Percent of support reach length that is dry (PCT_DRS)
If PCT_DR< 1%, score=1. If PCT-DR≥ 20%, then score=0.
Log_{10}[(mean wetted width x mean thalweg depth)+0.001] (LXWXD)
If LXWXD> 1, score=1. If LXWXD ≤ -1.4, then score=0
Residual pool depth (RP100)
If RP100 ≥20, then score=1. If RP100 ≤ 0, then score=0
Mean wetted width
If XWIDTH ≥ 6, then score=1. If XWIDTH=0, then score=0
HABITAT VOLUME INDEX=(PCT_DR score + LXWXD score +RP100 score + XWIDTH score)/4
PLOT NUMBER OF FISH COLLECTED (TOTLNIND) VS. HABITAT VOLUME INDEX (QVOLX)
Value for QVOLX below which most sites have no fish=0.43
PLOT HABITAT VOLUME INDEX VS. WATERSHED AREA (WSAREA_KM2)
QVOLX=0.42 corresponds to a watershed area of ~ 2 km^2

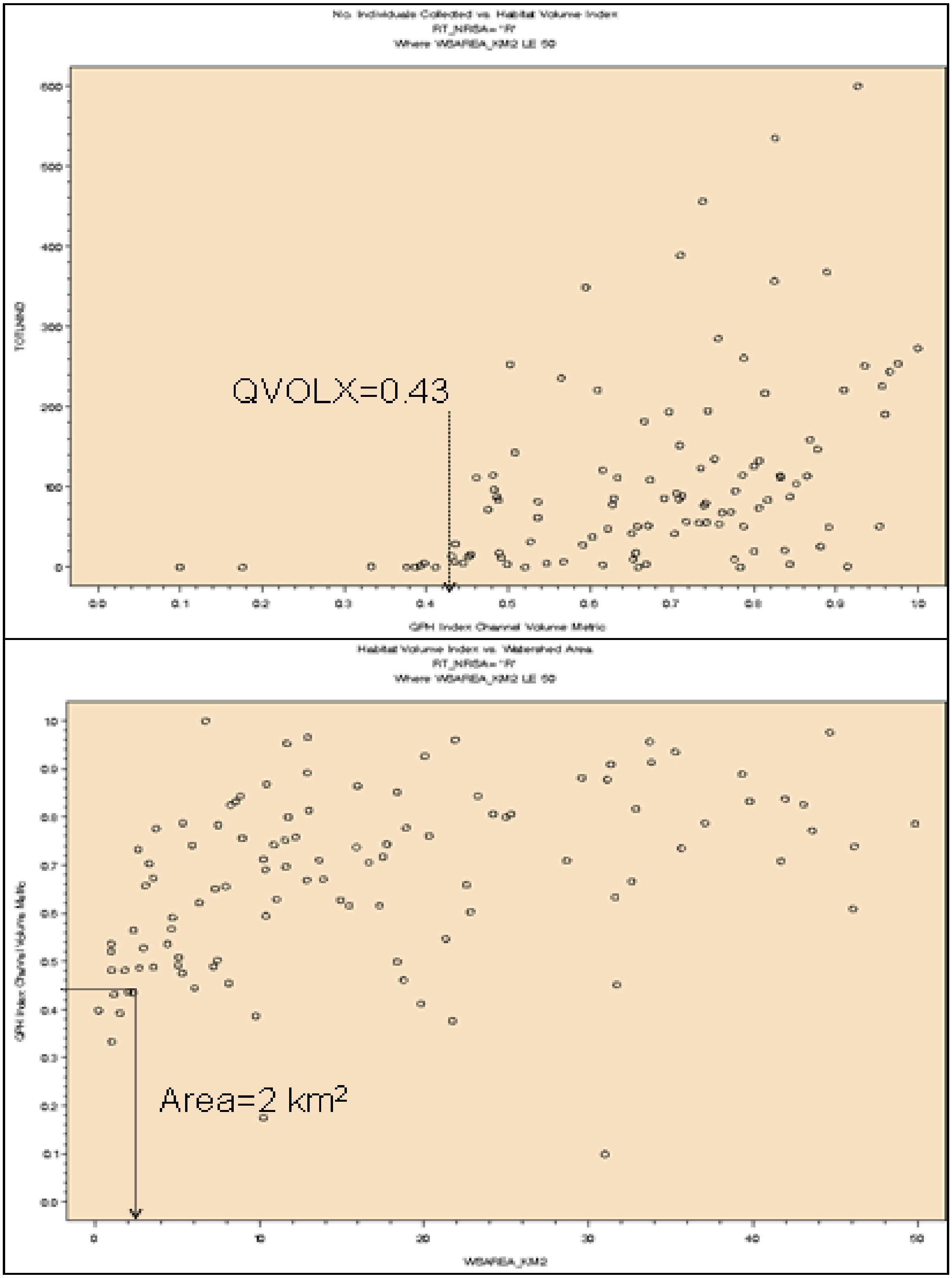


Figure D-3. Relationship between small watershed size, reduced habitat volume, and number of fish collected. Fish are not likely to be found in streams with a watershed area of <2 km^2.

4.4 FMMI performance

We evaluated the performance of the regional FMMIs in several ways (Table D-4). Comparing the FMMI scores from set of validation reference sites to those of the set of reference sites used for random forest modeling confirmed that the models were behaving as anticipated. For all three regional FMMIs, the mean values of the validation sites and sites used in modeling were not significantly different based on a two-sample *t*-test assuming unequal variances).

We evaluated the responsiveness of the regional FMMIs by comparing FMMI scores of the set of reference sites to the set of more highly disturbed sites (Stoddard et al. 2008). Boxplots (Figure D-4) and two-sample *t* tests (assuming unequal variances) showed that all FMMIs were highly responsive, but the FMMI for the West region was somewhat less responsive than the other two FMMIs (Table D-4).

We estimated precision of the models by calculating the standard deviation of FMMI scores from all reference sites, after standardizing the scores to a mean of 0. The FMMIs all appear to be very precise, with standard deviation value near 0.1 (Table D-4). These values are comparable (or better) than many predictive models of taxa loss (Hawkins et al. 2010a).

We evaluated the reproducibility of the regional FMMIs using a set of sites that were visited at least twice during the course of the NRSA project, typically two times in a single year (Kaufmann et al. 1999, Stoddard et al. 2008). We used a general linear model (PROC GLM, SAS v. 9.12) to obtain estimates of among-site and within-site (from repeat visits) variability. PROC GLM was used because of the highly unbalanced design (only a small subset of sites had repeat visits). We used a nested model (sites within year) where both site and year were random effects. We estimated reproducibility by deriving a "signal:noise" (S/N) ratio as $(F - 1)/c$, where F is the F-statistic from the ANOVA, and c is a coefficient in the equation used to estimate the expected mean square. If all sites had repeat visits, c would equal 2 (Kaufmann et al. 1999). If no sites had repeat visits, c would equal 1. For the Eastern Highlands, c = 1.1326, while for the Plains and Lowlands c = 1.0745 and for the West c = 1.0753. Values of S/N suggest the regional MMIs are reproducible, with values between 4 (West) and 8 (Plains and Lowlands; Table D-4).

Table D-4. Performance statistics for the three regional FMMIs.

Performance Characteristic	Eastern Highlands FMMI	Plains and Lowlands FMMI	West FMMI
Validation reference sites vs. reference sites used in metric modeling	t=0.13	t=1.19	t=-0.14
Reference sites vs. more highly disturbed sites	t=18.2	t=17.2	t=11.3
Model precision (	0.091	0.120	0.075
Reproducibility (Signal:Noise)	5.2	8.0	4.1

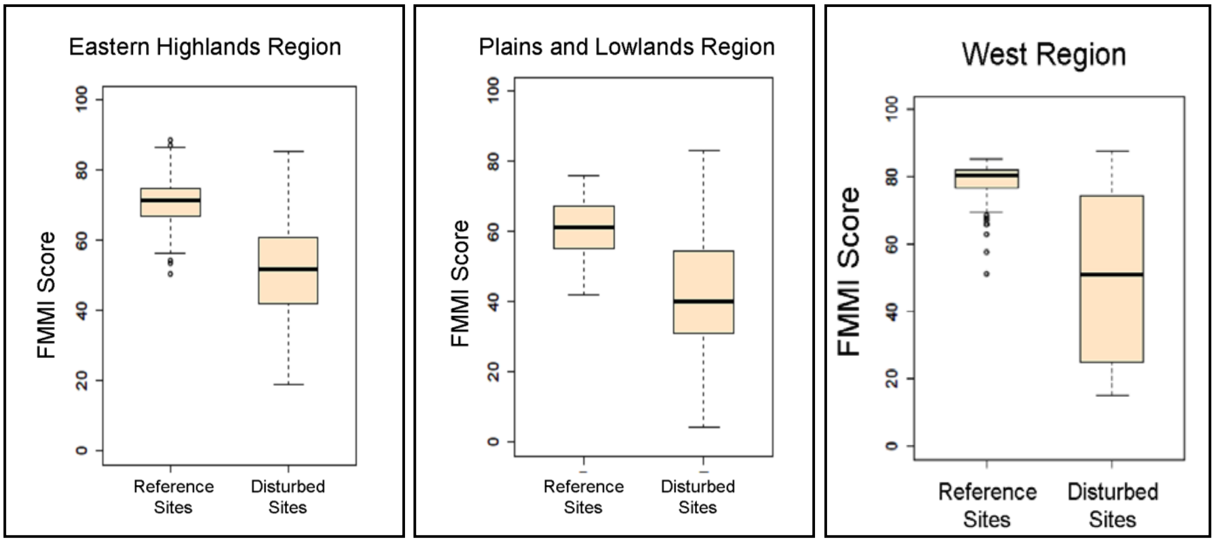


Figure D-4. Boxplots comparing FMMI scores of reference sites to more highly disturbed sites.

We felt it important to examine the performance of the component metrics across the range of stream sizes sampled for the NRSA. The potential exists for bias in the FMMI due to different fish species pools being available for larger rivers versus smaller streams. Differences across the size range might also result from the different sampling protocols that were used (wadeable, large wadeable, and boatable). We used the set of reference sites to examine patters in metric response values across Strahler order categories. The distribution of metric response values and FMMI scores among stream order classes does not indicate a bias due to either stream size or sampling method (Figures D-5 through D-7).

We looked at the distribution of FMMI scores across Strahler order categories as well. We did not expect to see much similarity, as disturbance is typically confounded with stream size. This pattern is evident in the Eastern Highlands and the West (Figure D-8).

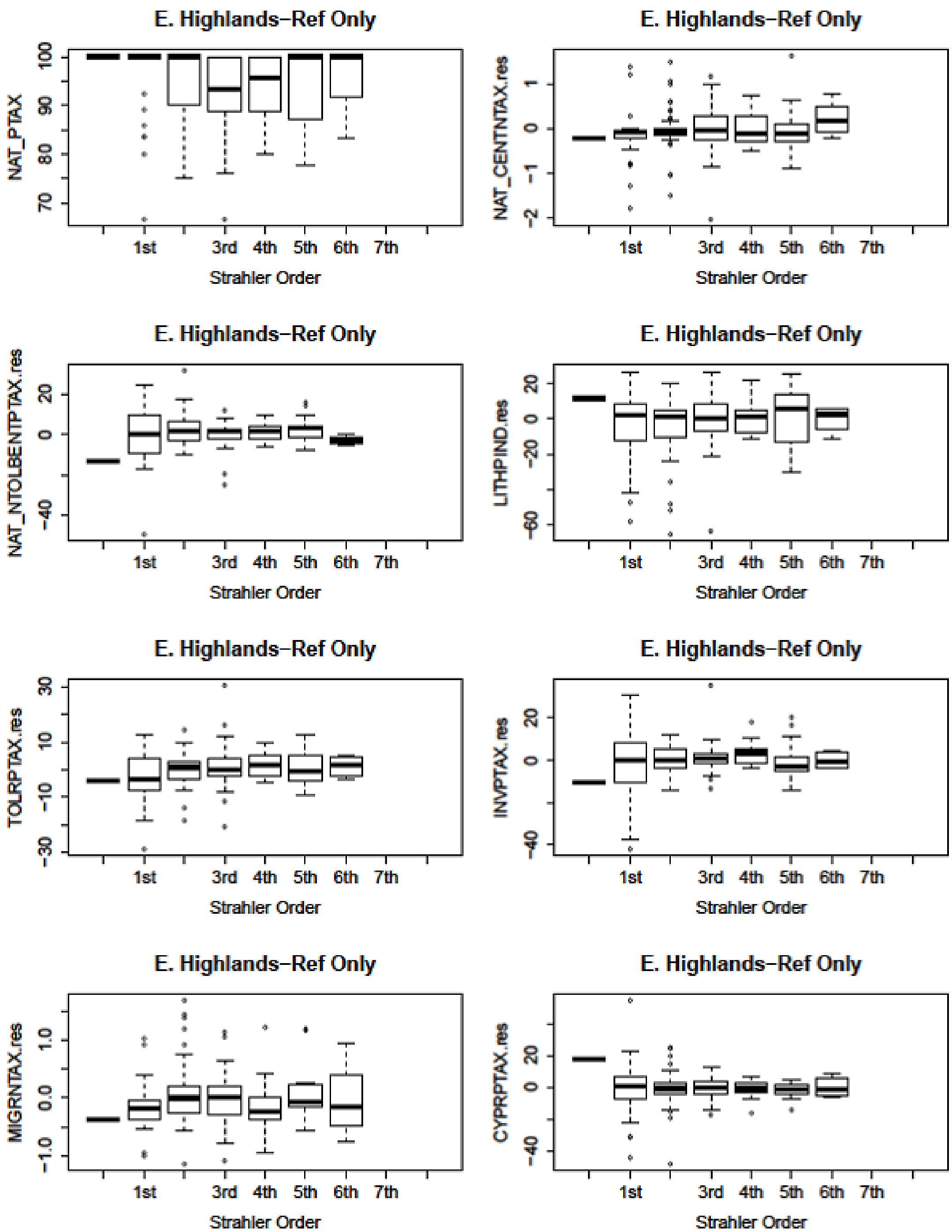


Figure D-5. Component metrics of the Eastern Highlands FMMI versus Strahler Order category,

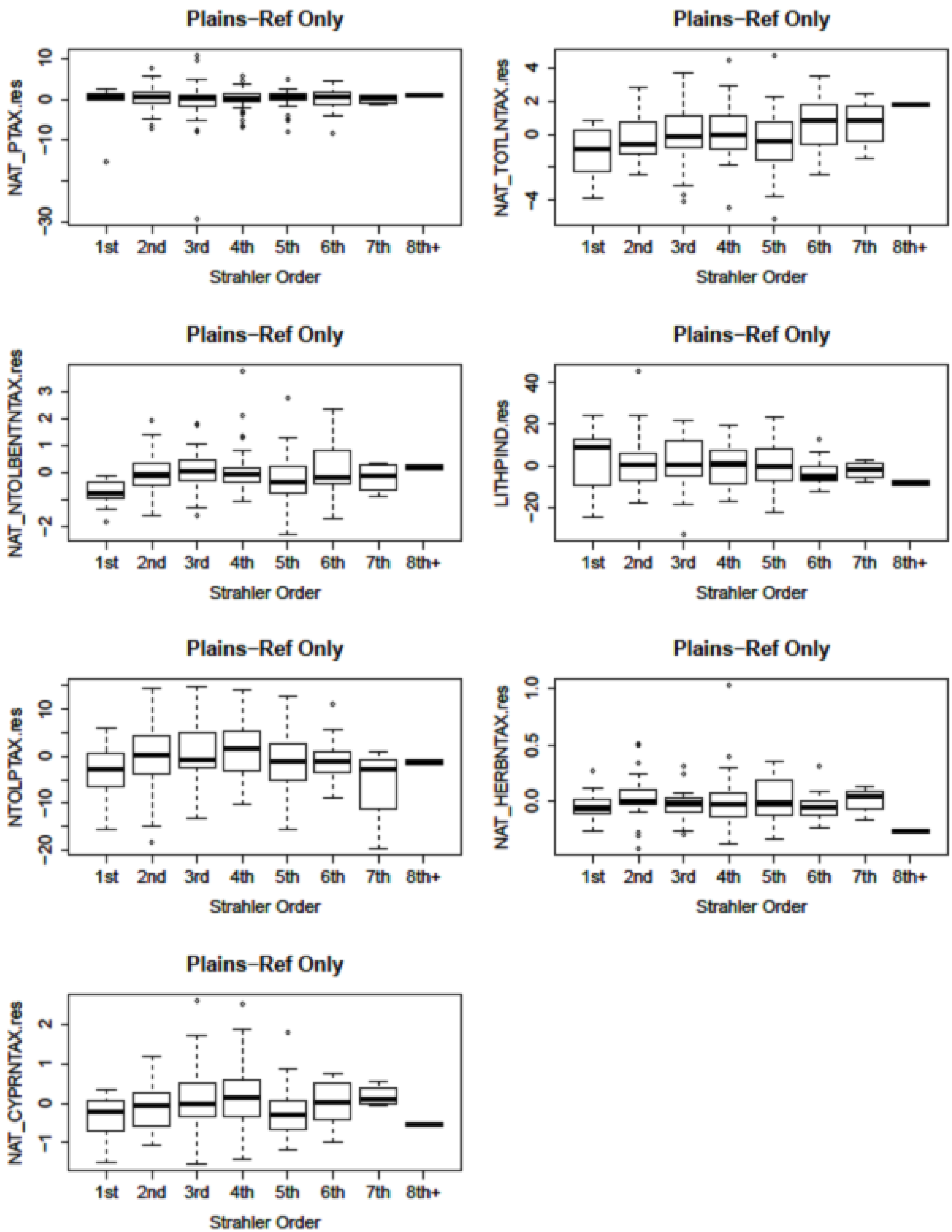


Figure D-6. Component metrics of the Plains and Lowlands FMMI versus Strahler Order category,

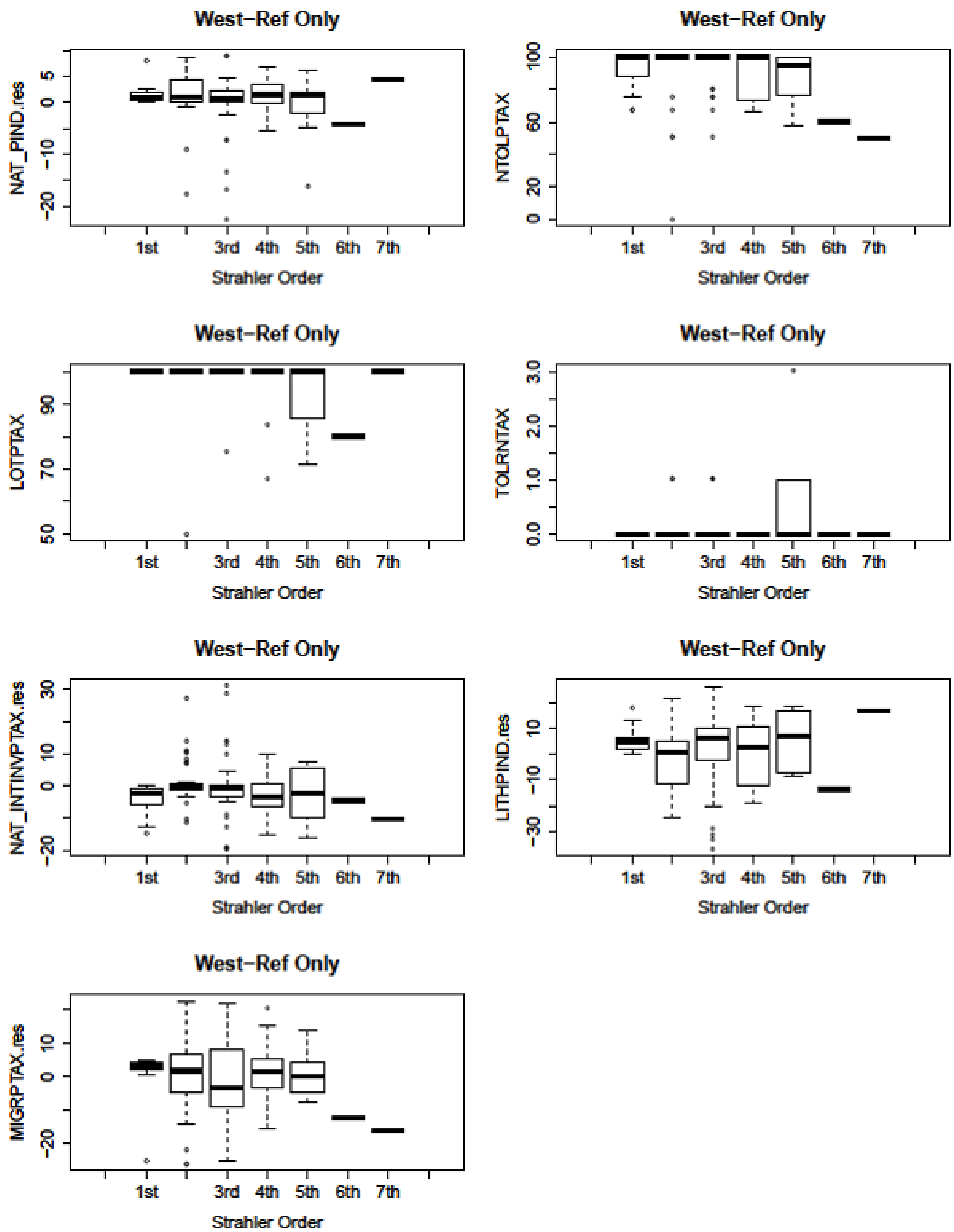


Figure D-7. Component metrics of the West region FMMI versus Strahler Order category,

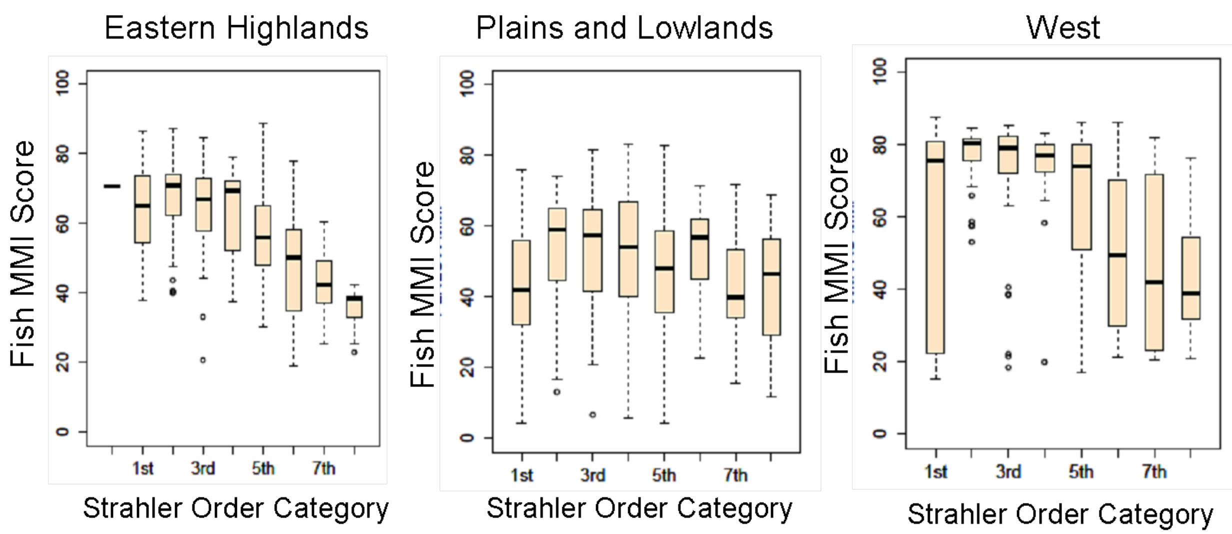


Figure D-8. FMMI scores versus Strahler order category.

4.5 Thresholds for assigning ecological condition

For the NRSA, ecological condition is based on the deviation from least-disturbed condition (Stoddard et al. 2006, Hawkins et al. 2010b). Within each of the three climatic regions, thresholds for defining "good" condition (similar to least-disturbed) and "Poor" condition (substantially different from least-disturbed) are based on the distribution of FMMI scores in reference sites in each of the nine aggregated ecoregions (Figure D-2). The threshold for "good" condition is equal to the 25th percentile of the distribution of reference sites within an aggregated ecoregion. The threshold for "poor" condition is set as being below the 5th percentile of the distribution of reference sites within an aggregated ecoregion.

Table D-5 presents the threshold values for the regional FMMIs. Sites with scores between the two threshold values were assigned a condition class of "fair" (indeterminate). In general, threshold values within a climatic regions differ from each other to prevent combining aggregated ecoregions to reduce the number of different thresholds being used. Three aggregated ecoregions contain fewer than 30 reference sites, so the threshold values (particularly those for fair/poor) are based on very few sites. In three regions, the difference between the two thresholds is less than 5 (Xeric = 2.9).

Table D-5. Thresholds for assigning ecological condition based on the distribution of regional FMMI scores in reference sites. Aggregated ecoregions are shown in Figures D-1 and D-2. Sample sizes are in parentheses.

Aggregated Ecoregion	Good/Fair (25th percentile)	Fair/Poor (5th percentile)
Eastern Highlands		
Northern Appalachians (60)	65.4	60.2
Southern Appalachian (94)	66.1	55.4
Plains and Lowlands		
Coastal Plains (39)	55.3	46.9
Northern Plains (42)	54.6	48.3
Southern Plains (43)	54.2	49.2
Temperate Plains (28)	60.3	52
Upper Midwest (28)	54.9	49.2
West		
Western Mountains (70)	67.5	57.6
Xeric West (25)	64.8	61.9

4.6 Literature cited

Barbour, M. T., J. Gerritsen, B. D. Snyder, and J. B. Stribling. 1999. Rapid bioassessment protocols for use in streams and wadeable rivers: periphyton, benthic macroinvertebrates, and fish. EPA 841/B-99/002, Office of Water. US Environmental Protection Agency, Washington, DC.

Blocksom, K. A. 2003. A performance comparison of metric scoring methods for a multimetric index for Mid-Atlantic Highlands streams. Environmental Management 31:0670-0682.

Bramblett, R. G., T. R. Johnson, A. V. Zale, and D. G. Heggem. 2005. Development and evaluation of a fish assemblage index of biotic integrity for northwestern Great Plains streams. Transactions of the American Fisheries Society 134:624-640.

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

Emery, E. B., T. P. Simon, F. H. McCormick, P. L. Angermeier, J. E. DeShon, C. O. Yoder, R. E. Sanders, W. D. Pearson, G. D. Hickman, R. J. Reash, and J. A. Thomas. 2003. Development of a multimetric index for assessing the biological condition of the Ohio River. Transactions of the American Fisheries Society 132:791-808.

EPA (United States Environmental Protection Agency). 2009. National Rivers and Streams Assessment: Field Operations Manual. EPA 841/B-04/004, Office of Water and Office of Environmental Information, US Environmental Protection Agency, Washington, DC.

Hawkins, C. P. 2006. Quantifying biological integrity by taxonomic completeness: its utility in regional and global assessments. Ecological Applications 16:1277-1294.

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

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

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

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

Karr, J. R. 1999. Defining and measuring river health. Freshwater Biology 41:221-234.

Karr, J. R. and E. W. Chu. 2000. Sustaining living rivers. Hydrobiologia 422/423:1-14.

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

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

Liaw, A. and M. Weiner. 2002. Classification and Regression by randomForest. R News 2:18-22.

Lyons, J., R. R. Piette, and K. W. Niermeyer. 2001. Development, validation, and application of a fish-based index of biotic integrity for Wisconsin's large warmwater rivers. Transactions of the American Fisheries Society 130:1077-1094.

McCormick, F. H., R. M. Hughes, P. R. Kaufmann, D. V. Peck, J. L. Stoddard, and A. T. Herlihy. 2001. Development of an index of biotic integrity for the Mid-Atlantic Highlands region. Transactions of the American Fisheries Society 130:857-877.

Meador, M. R. and D. M. Carlisle. 2009. Predictive Models for Fish Assemblages in Eastern U.S. Streams: Implications for Assessing Biodiversity. Transactions of the American Fisheries Society 138:725-740.

Mebane, C. A., T. R. Maret, and R. M. Hughes. 2003. An Index of Biological Integrity (IBI) for Pacific Northwest Rivers. Transactions of the American Fisheries Society 132:239-261.

Nelson, J. S., E. J. Crossman, H. Espinosa-Pérez, L. T. Findley, C. R. Gilbert, R. K. Lea, and J. D. Williams. 2004. Common and Scientific Names of Fishes from the United States Canada and Mexico. Sixth edition. Special Publication 29, American Fisheries Society, Bethesda, Maryland.

Page, L. M. and B. M. Burr. 1991. A field guide to freshwater fishes of North America north of Mexico. Houghton Mifflin, Boston, Massachusetts.

Pearson, M. S., T. R. Angradi, D. W. Bolgrien, T. M. Jicha, D. L. Taylor, M. F. Moffett, and B. H. Hill. 2011. Multimetric Fish Indices for Midcontinent (USA) Great Rivers. Transactions of the American Fisheries Society 140:1547-1564.

Pont, D., R. M. Hughes, T. R. Whittier, and S. Schmutz. 2009. A Predictive Index of Biotic Integrity Model for Aquatic-Vertebrate Assemblages of Western U.S. Streams. Transactions of the American Fisheries Society 138:292-305.

Pont, D., B. Hugueny, and C. Rogers. 2007. Development of a fish-based index for the assessment of river health in Europe: the European Fish Index. Fisheries Management and Ecology 14:427-439.

R Development Core Team. 2011. R: A language and environment for statistical computing R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Roset, N., G. Grenouillet, D. Goffaux, D. Pont, and P. Kestemont. 2007. A review of existing fish assemblage indicators and methodologies. Fisheries Management and Ecology 14:393-405.

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

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

Whittier, T. R., R. M. Hughes, J. L. Stoddard, G. A. Lomnicky, D. V. Peck, and A. T. Herlihy. 2007. A structured approach for developing indices of biotic integrity--three examples from western streams and rivers in the USA. Transactions of the American Fisheries Society 136:718-735.

Periphyton Assemblage Analysis for the NRSA 2008–2009

5.1 Background information

Many countries manage aquatic resources to protect or restore the structure and function that is characteristic of ecosystems with minimal disturbance by human activities. Multimetric indices of biological condition (MMIs) are commonly used in assessments of aquatic resource condition. As ecological assessments become more common and large scale, accounting for natural variation among aquatic resources has been a challenge for accurate assessment. Ecoregions account for some natural variation in climate, geology, hydrology, and soils among sites, but significant natural variation in size and slope as well as local hydrology and soils can occur within a region and affect the expected characteristics of fish, macroinvertebrate, and algal assemblages at a site.

Site-specific predictions of the expected characteristics of a site can be modeled and enable assessment of ecological condition as the deviation between observed and expected condition. Expected condition can be minimally disturbed condition or best available condition (*sensu* Stoddard et al. 2006), or desired condition (*sensu* Stevenson et al. 2004). Predictive models for expected condition can be determined using data from all sites in a regional assessment (Seelbach et al. 2002, Baker et al. 2005, and Riseng et al. 2010) or from selected sites that are assumed to meet management goals, such as reference sites (Clarke et al. 1996, Hawkins et al. 2000). In most cases, predictive models have been developed for metric of ecological condition, but recently Hawkins et al. (2010) developed invertebrate MMIs of biological condition for a region in the western U.S. using predictive models for each metric to account for natural environmental variability. These models use any metric, not just the number of expected taxa that traditionally occur at reference sites as in RIVPACS. Similar modeling of expected minimally disturbed condition has been routine with fish assemblages because of the great effects of region, water temperature, and stream size on fish metrics.

Algae are commonly used in ecological assessments of streams, lakes, wetlands, and coastal zones around the world, with periphyton being the most common algal assemblage in streams and wetlands. Surface sediment diatoms and phytoplankton are used in the National Lakes Assessment. Perphyton assemblages are used in the National Rivers and Streams Assessment and the National Wetland Condition Assessment. Metrics and MMIs of periphyton condition have been developed and tested for applications at regional scales, using complex metrics that limit detailed understanding of important elements of biological condition, and using diatoms alone versus the entire diatom and non-diatom periphyton assemblage.

For the NRSA, we developed and compared MMIs of biological condition using periphyton assemblages in rivers and streams of the United States. Metrics were calculated based on literature characteristics for diatom taxa and newly calculated characteristics for non-diatom algal taxa. Periphyton MMIs were compared to determine whether performance was greater if MMIs included non-diatom algal metrics as well as diatom metrics; for modeled rather than traditional MMIs, and for ecoregion-specific MMIs rather than one national MMI.

5.2 Methods

5.2.1 *Taxonomic analysis of periphyton samples*

Periphyton in 2,580 samples from streams and rivers of the United States were identified and counted for the EPA's National Rivers and Streams Assessment (NRSA). Samples were collected from random locations along 11 transects and combined into a composite sample. These samples came from 2,099 different rivers and streams, with 1,901 samples from sites designated as probability sites for the NRSA.

Periphyton samples were analyzed with separate steps for all algae and diatoms. In all algal counts, non-diatom algae were identified to the lowest taxonomic level (diatoms were not identified to lower taxonomic levels) and both cells and natural units counted, live diatoms were counted, and dead diatoms (those without protoplasm in frustules) were counted. Natural units were cells for unicellular algae and a filament or a colony for algae that occur in multicellular growth forms. Count and identifications of all algae were conducted in a Palmer-Maloney counting chamber at 400X and continued until 300 natural units had been identified or observed. Diatoms were acid-cleaned and mounted in NAPHAX on microscope slides before 500 valves were identified to the lowest taxonomic level. The lowest taxonomic level was almost always species or variety for diatoms and was commonly species for all other algae, except some green algae that require reproductive cells to identify species.

Counts of all algae and live diatoms were aggregated into a single count by multiplying number of live diatom cells by the proportional relative abundance of individual diatom taxa, and replacing the live diatom record in counts of all algae with the records of the individual diatom taxa. Counts of live and dead diatoms were recorded in a separate table from counts of all algae.

Taxonomic consistency was increased by exchange of images of all taxa with special attention to problem taxa. A workshop was conducted mid-way through the period when counts were being done so taxonomists had experience with the diversity of taxa in samples, but had not finished counts. This enabled refinements of taxonomy after the workshop with conference calls and ongoing exchanges of images. A taxonomic harmonization table was constructed by listing all names used by taxonomists in the project, which we referred to as NRSA taxon name. NRSA taxa names were the finest level of taxonomy reported in the project and allowed taxonomists to distinguish between taxa as well as they could. In some cases, some taxonomists did not distinguish between taxa in the same way. To account for situations when one taxonomist distinguished between a set of taxa and another did not, we retained the finer distinction in NRSA taxon names, but then assigned the multiple distinguished taxa to a single name that was more widely used by other taxonomists. This kind of lumping taxa together and assigning coarser and coarser levels of operational taxonomic units was done twice, resulting in lump level 1 names and lump level 2 names. During the taxa harmonization process, the goal of lump level 1 (LL1) names were to provide a set of names that would be consistent for ecological analysis. Further lumping than lump level 2 was considered unnecessary by taxonomists and could be accomplished in data analysis by using genera. Since count data were reported at the

NRSA taxon name level, count data were aggregated by the LL1 name for metric and MMI calculation.

5.2.2 *Candidate metric calculation*

Candidate periphyton metrics were calculated using tables of all algae in counts, only diatoms, and only non-diatom algae. Three categories of metrics were calculated, each representing different dimensions of biological condition: taxonomic composition, ecological function, and diversity. Candidate metrics for taxonomic composition of periphyton were calculated using relative abundance of taxa that were sensitive and tolerant to human disturbance (sens-tol metrics) and relative abundances of genus and higher levels of taxa (family to phylum). Candidate metrics for ecological function were calculated with relative abundances of individuals with different ecomorphological traits. Shannon diversity, number of taxa observed in counts, and number of taxa observed in counts per individual (S/N) were measures of diversity. Algal biomass was a fourth metric category that was considered, but measures of chlorophyll a, ash-free dry mass, and filamentous algal cover were not made for a large proportion of sites.

Sens-tol periphyton metrics were calculated using traits for diatom taxa that were developed using data in previous projects and newly determined traits for non-diatom taxa using NRSA data. We used existing traits of diatoms to provide complete independence and test reliability of using traits for the literature, versus recalculating diatom traits with NRSA data and then using that to assess biological condition. We used new traits for non-diatom taxa because sensitivity and tolerance of many taxa of non-diatom algae to human disturbance has not been determined in previous studies. Traits were selected based on attributes of biological condition described in Davies and Jackson (2006), which can be indicated by the relative abundances of taxa that are sensitive and tolerant to a generalized stressor gradient of human disturbance. Characterizations of biological condition in the National Aquatic Resources Surveys emphasize independent assessment of taxa characteristic of minimally disturbed and highly disturbed conditions. Classification of sites as minimally disturbed (reference) sites, moderately disturbed, and highly disturbed is described below. We intentionally excluded metrics calculated with taxa traits based on sensitivity and tolerance to individual pollutants (e.g., nutrients, pH, or conductivity), because they have been highly correlated with metrics for the generalized stressor gradient of human disturbance and there is some concern that that traits are uniquely indicative of the pollutant for which they were calculated (Stevenson et al. 2008, Stevenson et al. submitted). In addition, the processes of characterizing biological condition and then identifying pollutants affecting biological condition should be as independent as possible (Stevenson 2006).

Traits have been assigned to algal taxa based on: 1) characteristics of sites in which they have highest relative abundances; 2) morphological attributes that are assumed to confer sensitivity or tolerance to human disturbance or to have ecologically significant functions, or 3) higher-level taxonomic groupings (genus-order level taxonomy) in which we assume evolutionary constraints limit variability of traits and restrict taxa to one habitat or another. Characteristics of sites in which taxa have their highest relative abundances can be determined by weighted average regression, linear regression, or indicator species analysis. Weighted average regression enables determination of environmental optima along a gradient of environmental conditions, such as optimal levels of pH, total phosphorus, or human disturbance. Linear regression has been used to determine whether taxa are characteristically found at one or the other end of an environmental gradient, i.e. low or high pH, low or high total phosphorus, or low or high human disturbance. Indicator species analysis determines whether taxa are characteristically found in a group of sites, whether at the low or high end of an environmental gradient, or potentially in the middle of that gradient. When taxa are characteristically found in specific environmental conditions, they are assumed to have morphological and physiological traits that enable their colonization, reproduction, competition, and persistence in those habitats.

Ecomorphological traits, traits assumed to confer sensitivity or tolerance to human disturbance or to have ecologically significant functions, were assigned to NRSA periphyton taxa based on the taxonomic and ecological literature or best professional judgment. Traits were assigned at the genus and higher taxonomic levels if taxa were filamentous or colonial, had heterocytes, were motile in sediments by means of raphe, flagellum, or gliding through sheaths, were highly adnate (grazer resistant) or stalked (grazer sensitive), or were highly characteristic of planktonic, benthic, epiphytic, and epipsammic (small and highly adnate to enable living on sand) habitats.

Taxonomic metrics were calculated as the number of cells or lowest-level taxa in the higher taxonomic levels, including diatom genera and families, orders, classes, and phyla of all algae.

Sensitive and tolerant classifications of non-diatom algal taxa were determined using indicator species analysis (Dufrene and Legendre 1997) to determine if species were characteristically found in reference or highly disturbed sites. Indicator species analysis was conducted for any taxon (phyla, classes, orders, families, genera, and LL1 names) of non-diatom algae if that taxon was observed in 30 or more samples. Indicator species analyses were calculated separately using proportions of natural units, biovolume, and cells of taxa in samples. For many taxa, sensitive or tolerance classifications at more than one taxonomic level were observed. The sensitive or tolerant characterizations of all taxa were compiled so characterizations at the phylum, class, order, family, genus, and LL1 name, if available, were recorded. In many cases, characterizations at lower taxonomic levels were not available because those taxa were not observed in 30 or more samples.

Diatom traits were selected for use in metrics if they could be attributed to a multistressor human disturbance gradient. Many traits of diatom taxa characterize their sensitivity and tolerance to specific physical and chemical conditions of habitats, such as pH and total phosphorus, versus a more general sensitivity or tolerance to human disturbance. Although diatom metrics for independent stressors can be highly precise and related to human disturbance, our goal was to separate assessment of biological condition and diagnosis of stressors affecting biological condition, which we plan to do in future research. Thus, a limited number of diatom traits characterizations met our criteria. The pollution tolerance index traits from Bahls (1993) and Lange-Bertalot (1979) assign taxa to sensitive and tolerant classifications based on variations in taxa relative abundances at sites with high or low levels of pollution. We used the Bahls and Lange-Bertalot pollution tolerant traits reported in Porter et al. (2008) because the taxonomy from the earlier literature had been updated in Porter et al. (2008). Stevenson et al. (2008a) used indicator species analysis and linear regression to characterize taxa as sensitive and tolerant to human disturbance using data from the western United States.

Candidate metrics based on taxa sensitivity and tolerance to human disturbance, ecomorphological traits, and higher-level taxonomic groups were calculated as the number of taxa in samples with a specific trait, the proportion of taxa in samples with a specific trait, and the proportion of individuals in samples with a specific trait. As a result, three candidate metrics were calculated for each trait and over 100 metrics were calculated. Weighted average metrics (in which different indicator values are assigned to a taxon for a trait and indices are calculated as the sum of products of indicator values times proportional relative abundances of taxa in samples) were not calculated because changes in sensitive and tolerant taxa cannot be independently evaluated, which is deemed important in assessment of biological condition (Davies and Jackson 2006, Stevenson 2006).

Sens-tol metrics using traits from Bahls (1993), Lange-Bertalot (1979), and Stevenson et al. (2008a) and all diatom taxa in samples were calculated as the relative abundance of the sensitive or tolerant units in a sample divided by the sum of sensitive and tolerant units in a sample. Thus, sens-tol metrics measured proportional abundance of either sensitive or tolerant taxa or individuals in relation to the number of taxa or individuals in a sample for which sensitivity or tolerance was known. In an ideal situation, we would know the sensitivity or tolerance of all taxa or individuals; but since we used existing trait information from the literature, a sensitivity and tolerance to human disturbance was not known for many taxa. Therefore, either sensitive or tolerant metrics using traits from Lange-Bertalot, Bahls and Stevenson et al. were considered for use in MMIs because the sensitive metric was the inverse of the tolerant metric. In contrast, the proportion of sensitive and tolerant taxa or individuals of diatom orders was determined as the simple sum of all taxa or individuals with those traits and the total number of taxa or individuals in those orders. In the latter case, both sensitive and tolerant metrics for taxa or individuals of diatom orders could be used to because sufficient numbers of them were unknown and sensitive metrics were not necessarily inversely and highly correlated with tolerant metrics.

Diversity metrics were used in our assessment, despite the potential for misinterpretation of their meaning. The diversity metrics used for periphyton were expected to increase with human disturbance because: nutrient pollution is a prominent effect of human disturbance of watersheds; low nutrient concentrations characteristic of minimally disturbed conditions are so low that they limit the species membership and constrain species richness; thus moderate levels of nutrient pollution probably increases both the numbers of species in habitats and evenness of their abundances. Thus, the increase in diversity along human disturbance gradients is a result of the invasion of taxa that require high nutrient concentrations. In contrast, we expect to observe decreases in indicators of sensitive reference taxa, such as proportion of sensitive taxa in samples. Thus, the effects of nutrient pollution on biological condition as indicated by presence of sensitive native taxa (*sensu* Davies and Jackson 2006) is negative, whereas invasion of species tolerant to nutrient pollution (actually requiring high nutrients for successful colonization and reproduction) indicated poor biological condition. These concepts are explained in more detail in Stevenson et al. (2008b). Thus, the diversity metrics calculated for periphyton are expected to increase with human disturbance of watersheds such that high diversity metrics indicate fair or poor biological condition.

Candidate metrics were recalculated using random forest models and multiple linear regression models to account for natural variation among sites and for later use in modeled MMIs. The recalculation or modeled metric value was determined as the difference between observed and expected values of a metric at a site. Modeled metric values greater and less than zero were, respectively, greater and less than expected metric values if the site was minimally disturbed. To develop the random forest and multiple linear regression models of expected condition, only counts from two-thirds of the reference sites were used. One-third of the reference sites were randomly selected and used later to validate the modeled MMIs. A set of climate, geological, and landscape variables that are affected little by human activities were selected to be independent variables in random forest and linear regression models of the expected metric value for minimally disturbed conditions at a site. The set of climate, geological, and landscape variables included latitude and longitude of the site as well as the following watershed characteristics: area, mean elevation, erodible soils, slope, maximum and mean temperatures, annual precipitation, watershed elongation, runoff potential, water capacity of soils, organic matter in soils, permeability of soils, depth of soils to bedrock, and a hydrologic stability factor. Modeled metrics were calculated separately for reference sites across the country and for each of three large ecoregions (Eastern Highlands, Lowland Plains, Western Nountains) across the United States to compare MMIs at the national and ecoregion scale.

5.2.3 *Metric selection and MMI calculation and evaluation*

Sensitivity of metrics to human disturbance, commonness, and independence of candidate metrics was used to select metrics for MMIs. Sensitivity to human disturbance was measured with t-statistics comparing central tendency and variation between reference and highly disturbed sites. Commonness was determined by the percent of sites with taxa present for calculation of metric values, with 75% as a criterion for acceptability. Metric independence was shown if candidate metrics were not highly correlated ($r^2>0.64$) among reference sites, indicating they were relatively independent characterizations of biological condition. In the some cases, we allowed covarying metrics in an MMI if they were in the same metric category and if they contributed to equal weights of metric categories in the MMI. Priority was also given to candidate metrics with consistently high performance across the ecoregions so the same metrics could be used in MMIs for all ecoregions.

Metric selection was conducted in a two-stage process. First, sensitivity and commonness were used to reduce the number of metrics within each metric category before metrics were modeled and evaluated for inclusion in the MMI. Then, final metrics for MMIs were chosen after sensitivity of modeled metrics was determined.

Separate MMIs were calculated for the national dataset and for each ecoregion. Metrics, whether modeled or traditional, were recalculated to a standard range of 0 to 10 by establishing the range of metrics between the 5th and 95th percentile of metric values, subtracting the 5th percentile from all metric values, dividing that difference by the range between the 5th and 95th percentiles of each metric's values, and then setting any values below zero to zero and above 10 to 10. If metric values were lower in reference than disturbed sites, the value of metrics was reversed by subtracting them from 10 to make high metric values indicate high levels of biological condition. The MMIs were then calculated as the sum of all metric values and rescaling that sum so the MMIs would have a maximum of 100 and minimum of zero.

Modeled MMIs were compared with validation measures of modeled MMIs, variation in modeled values at reference sites, and by t-statistics comparing modeled MMIs at reference sites and highly disturbed sites. Modeled MMIs were validated by comparing the median and range of modeled MMI values in the 66.7% of reference sites used to develop the models for expected metrics values and the 33.3% of references sites not included in metric model development. MMI performance was related to low variation among reference sites and high t-statistics between reference and highly disturbed sites. MMI performance was measured by the variation in response to the human disturbance gradient, indicated by either agricultural, rowcrop, or urban land use in watersheds or total phosphorus concentration. These land use measurements and total phosphorus concentration have been highly correlated to indices of periphyton biological condition in past research. MMIs had to clearly distinguish between reference and highly disturbed sites to be used in the NRSA. The decision of whether traditional or modeled MMIs should be used and whether MMIs for the whole nation or by ecoregion should be used was based on MMI performance and satisfactory model validation.

Biological condition of the streams and rivers of the United States was determined by comparing MMI values at all probabilistic sites in the NRSA to the 75th and 5th percentiles of MMIs at reference conditions at both the national and ecoregion scales.

5.3 Results

5.3.1 *Non-diatom trait characterization*

Many non-diatom periphyton metrics responded sensitively to human disturbance gradients. Indicator species values indicated that red algae (Rhodophyta) were characteristically found at reference sites and euglenoids (Euglenophyta) were found in highly disturbed sites. Many other orders, families, and genera of green algae and cyanobacteria were also identified as indicator species of reference and highly disturbed conditions. Many green algal orders had significant indicator values for reference conditions, including the Desmidales, Zygnematales, and Ulotrichales. Indicator conditions for cyanobacteria (Cyanophyceae) differed by family and genus. Characteristically planktonic orders, such as the family Nostocaceae and the genus *Anabaena* or the family Microcystaceae and genus *Microcystis*, were characteristically found in highly disturbed conditions. Alternatively, the cyanobacteria family Rivulariaceae and the genus *Nostoc* within the family Nostocaceae were characteristically found in reference conditions. The cyanobacteria genera *Merismopedia*, *Spirulina*, and *Planktothrix* were characteristically found in highly disturbed condition, whereas *Homeothrix*, *Chroococcus*, *Microcoleus*, and *Woronichinia* were found in reference conditions. Similar genus and LL1 specific differences in reference or highly disturbed indicator conditions were observed for some green algae. In most cases, when genera and LL1 taxa were characteristically found in either reference or highly disturbed conditions, the higher-level taxa (families, orders, and classes) were characteristic of the same conditions, if the higher-level taxa were characteristically found in any habitat. But exceptions were noted for *Nostoc* and a *Pediastrum* species.

5.3.2 ***Metric selection***

Some metrics could not be calculated for sites because no taxa with trait characterizations for those metrics were observed in the sample for those sites. Many metrics based on diatom genera were observed at less than 75% of sites. Metrics based on non-diatom algae were measureable at less than 75% of the sites either because samples were composed entirely of diatoms or observed non-diatom algae were not characteristic of reference or highly disturbed sites. Non-diatom metrics were not considered further in assessment of metric because they could not be calculated at so many sites using the results of all algae counts of 300 natural units. Longer counts of non-diatom algae would likely provide sensitive and repeatable metrics. The level of effort associated with generating those metrics and the benefit-cost ratio compared to just using diatoms should be determined in future research.

Twelve sensitive-tolerant species composition metrics, five taxonomic metrics, 11 functional group metrics, 12 diversity metrics, and the live diatom metric were measured in over 90% of samples and were modeled using random forest and multiple linear regression (Table E-1). The sensitive-tolerant species composition metrics using all diatom taxa and for taxa in specific orders of diatoms had the highest absolute values of t-statistics for comparisons of reference and highly disturbed sites for unmodeled as well as both modeled forms of metrics (Table E-2). This held true for national-scale as well as ecoregion-specific t-statistics. r^2 values for random forest and multiple linear regression models for sensitive-tolerant species composition metrics were generally greater than a 0.10 guideline for application.

Taxonomic metrics had lower absolute values of t-statistics than the sensitive-tolerant species composition metrics. In addition, r^2 for random forest and multiple linear regression models of taxonomic metrics were more commonly lower than the 0.10 benchmark than for sensitive-tolerant species composition metrics. Proportion of *Achnanthidium* individuals and proportion of *Nitzschia* individuals had the highest absolute values of t-statistics of the taxonomic metrics.

Some functional metrics also had high absolute values of t-statistics, including proportion of planktonic cells of all algae, proportion of planktonic diatoms, and proportion of highly motile diatom tax and proportion of stalked diatoms. r^2 for random forest and multiple linear regression models of functional group metrics were more commonly lower than the 0.10 benchmark than for sensitive-tolerant species composition metrics.

Some diversity metrics had t-statistics in the acceptable range (≥3.0). In general diatom diversity metrics had higher absolute values of t-statistics than diversity metrics based on all algae. Diversity metrics increased with human disturbance as predicted.

Overall, multiple linear regression models explained more variation in metrics among reference sites than random forest models. However, the difference in variation explained (r^2 values) was relatively low for metrics that had the highest r^2 values and the highest t-statistics.

In general, unmodeled metrics had higher absolute values of t-statistics than metrics modeled with random forests, which in turn had higher absolute values of t-statistics than metrics modeled with multiple linear regression. As a result, metrics modeled with random forests were selected over metrics modeled by multiple linear regression.

Correlation among metrics was generally low and seldom exceeded an r^2 value of 0.64.

Metrics with the highest absolute values of t-statistics at the national scale typically had the highest absolute values of t-statistics at the ecoregion scale, also. Therefore, the same metrics were used for MMIs for the national scale and ecoregions, as well as for traditional MMIs and for modeled MMIs. The following metrics were selected for use in MMIs.

- Proportion of sensitive (diatom) taxa was selected because it had t-statistics almost as high as proportion of sensitive (diatom) individuals, it is an important characteristic of assemblages to record, and it maximize conceptual consistency with metrics based on proportion of sensitive and tolerant taxa of diatom orders.
- The proportion of sensitive Achnanthaceae taxa, the proportion of sensitive Diatomaceae taxa, the proportion of sensitive Naviculaceae taxa, as well as the proportion of tolerant Naviculaceae taxa and proportion of tolerant Nitzschiaceae taxa were selected for the MMIs because their t-statistics were usually nearly as high as the versions of these metrics calculated as proportion of individuals and changes in proportions of taxa in samples with specific attributes is more directly related to biological condition than changes in proportions of individuals.
- Because taxonomic metrics were very similar to the sensitive-tolerant species composition metrics based on diatom orders, we did not use any taxonomic metrics.
- The proportion of highly motile diatom individuals, stalked diatom individuals, and planktonic diatom individuals were selected as functional metrics because they had the highest absolute values of t-statistics.
- The number of diatom taxa per individual (S/N) diversity of diatoms, Shannon diversity of diatoms, and number of diatom families per individual were used as diversity metrics.
- Percent live diatoms was not used as a metric because it had relatively low t-statistics.

As a result of this metric selection process, six metrics based on the sensitive-tolerant species composition were used in MMIs with three functional group metrics and three diversity metrics. A couple of the diversity metrics were highly correlated, but they were used so that three diversity metrics could be used in the MMI and provide the same weight for the diversity category of metrics compared to functional metrics. Because of the relative independence of the sensitive-tolerant species composition metrics based either on all diatom taxa or by diatom order and the large number of these metrics having high t-statistics, we decided that this category of metrics should have twice as much weight as functional or diversity metrics.

5.3.3 *MMI comparison and evaluation*

The traditional and the modeled MMIs were significantly different between ($p<0.001$) reference and highly disturbed sites whether calculated at the national scale or for any specific ecoregion. Median MMIs for either reference sites or highly disturbed sites were not within the interquartile ranges of the other category of sites for any of the MMIs (Figure E-1). The difference in median MMIs between reference and highly disturbed sites was greater for traditional than modeled MMIs for the national and Lowland Plains models but not for the Eastern Highlands and Western Mountains models. In contrast, interquartile ranges of the traditional and modeled MMIs for reference and highly disturbed sites overlapped for the MMIs calculated with all sites for the nation and the Lowland Plains ecoregion, but interquartile ranges did not overlap for the Eastern Highlands and Western Mountains ecoregions. The variation in modeled MMIs among reference sites was very low in the Eastern Highlands and Western Mountains ecoregions compared to the modeled MMI at the national scale and for the Lowland Plains ecoregion as well as any traditional MMI. In summary, modeled and ecoregion-scale periphyton MMIs had lower variation in reference condition and greater discrimination between reference and trashed sites than traditional or national-scale MMIs.

Modeled MMIs were validated well in comparisons of the medians and ranges of MMIs calculated for the two-thirds of reference sites used to develop the models for metrics at reference sites and the one-third of reference sites not included in model development (Figure E-1). Median MMIs of calibration and validation reference sites were very close. Ranges for validation sites of modeled MMIs were moderately greater than for the calibration sites, but the modeled MMI ranges for validation sites were not biased high or low compared to calibration sites.

The proportions of sites classified as having good, fair, and poor biological condition using 25^{th} and 5^{th} percentiles of reference condition varied greatly depending upon whether traditional or modeled MMIs were used and whether one MMI was used for all sites or separate MMIs were used for ecoregions. The proportions of sites classified as fair and poor were greater for modeled MMIs than traditional MMIs and greater for ecoregion-scale than national-scale MMIs (Figure E-2). In some cases, these differences were great. For example, the number of all probabilistic sites classified as good would be approximately 60% using the traditional national-scale MMI and approximately 45% using the modeled national-scale MMI, but only approximately 10% of sites would be classified as good in the eastern highlands ecoregion using the modeled MMI for that region if the 25^{th} percentile of reference condition was used as the benchmark for that good status.

Both traditional and modeled MMIs were related to the percent of watershed altered by human activities ($p<0.001$), but the magnitude of response was greater for traditional MMIs than modeled MMIs, especially for the national and Lowland Plains models (Figure E-3, Table E-3). The amount of variation in MMIs explained by percent watershed disturbed was also greater for traditional than modeled MMIs. The amount of variation in MMIs explained by the log-transformed percent of watersheds altered by humans (+1 to account for zero percentages) ranged from 22% for the traditional MMI at the national scale to 2% for the model MMI for the lowland plains ecoregion. The proportional decrease in response magnitude and variation explained between traditional and modeled MMIs was greater for the national-scale analysis and for the Lowland Plains than for the Eastern Highlands and Western Mountains ecoregions. Conclusions were similar to those above from more detailed regression analysis of MMIs and human disturbance in watersheds when percent agriculture, rowcrop, and urban activities were distinguished. The later analyses are not presented in the report.

Both traditional and modeled MMIs were related to total phosphorus concentration in streams ($p<0.001$, Figure E-4). The magnitude of responses of MMIs were relatively similar among ecoregions compared to MMI responses to percent watershed altered. The magnitude of response to total phosphorus was greater for traditional MMIs than modeled MMIs. The amount of variation in MMIs explained by total phosphorus was also greater for traditional than modeled MMIs.

The percent of watersheds altered by humans was greater in the lowland plains than the eastern highlands and western mountains ecoregions (Figure E-5). Watershed alteration by humans was lowest in the western mountains. The proportion of agricultural land use used for rowcrops was greater in the lowland plains than the eastern highlands ecoregion.

Alteration of watersheds by human activities was related to natural environmental factors. For example, the percentages of agricultural and urban land use was related to elevation of watersheds, whether all sites or just reference sites were considered (Figure E-6).

5.4 Discussion

Ecoregion-specific modeled MMIs were selected to assess biological condition of periphyton in the National Rivers and Stream Assessment. The variation in MMI among reference sites was less for modeled than traditional MMIs. The discrimination of reference and trashed sites by MMIs was greater for modeled than traditional MMIs. Although the magnitude of response of modeled MMIs to percent watershed disturbed was not as great as traditional MMIs, this was attributed to covariation among natural ecological factors used in modeling expected values of metrics and where human activities occur. Thus, we often observed a lower absolute difference in metric values between reference and highly disturbed sites or along human disturbance gradients for modeled versus traditional MMIs, but our characterization of expected reference condition was more precise and our discrimination of reference and disturbed sites was greater for modeled versus traditional MMIs.

The lower response of modeled MMIs to human disturbance in the Lowland Plains ecoregion may be due to lower sensitivity of Lowland Plains streams to human disturbance or to greater human disturbance in watersheds of reference sites for the Lowland Plains versus the Eastern Highlands or Western Mountains ecoregion. Percent watershed alteration by humans for reference sites in the Lowland Plains was greater than in the Eastern Highlands and Western Mountains; but to determine whether Lowland Plains streams are more resistant to human disturbance, we will need to use different methods for characterizing biological condition and expected condition in minimally disturbed conditions to recalibrate assessment of biological condition of periphyton in streams. One approach for recalibrating assessments is to use land use explicitly in modeling of the expected minimally disturbed condition.

5.5 Literature cited

Bahls, L. L. 1993. Periphyton bioassessment methods for Montana streams. Montana Department of Health and Environmental Sciences, Helena, Montana, US.

Baker, E. A., K. E. Wehrly, P. W. Seelbach, L. Wang, M. J. Wiley, and T. Simon. 2005. A multimetric assessment of stream conditoin in the Northern Lakes and Forests ecoregion using spatially explicit statistical modeling and regional normalization. Transactions of the American Fisheries Society **134**:697-710.

Clarke, R. T., M. T. Furse, J. F. Wright, and D. Moss. 1996. Derivation of a biological quality index for river sites: comparison of the observed with the expected fauna. Journal of Applied Statistics **23**:311-332.

Davies, S. P., and S. K. Jackson. 2006. The biological condition gradient: a descriptive model for interpreting change in aquatic ecosystems. Ecological Applications **16**:1251-1266.

Dufrene, M., and P. Legendre. 1997. Species assemblages and indicator species: the need for a flexible asymmetrical approach. Ecological Monographs **67**:345-366.

Hawkins, C. P., R. H. Norris, J. N. Hogue, and J. W. Feminella. 2000. Development and Evaluation of Predictive Models for Measuring the Biological Integrity of Streams. Ecological Applications **10(5)**:1456-1477.

Hawkins, C. J., Y. Cao, and B. Rober. 2010. Method of predicting reference condition biota affects the performance and interpretation of ecological indices. Freshwater Biology **55**:1066-1085.

Lange-Bertalot, H. 1979. Pollution tolerance of diatoms as a criterion for water quality estimation. Nova Hedwigia **64**:285-304.

Porter, S. D., D. K. Mueller, N. E. Spahr, M. D. Munn, and N. M. Dubrovsky. 2008. Efficacy of algal metrics for assessing nutrient and organic enrichment in flowing waters. Freshwater Biology **53**:1036-1054.

Riseng, C. M., M. J. Wiley, P. W. Seelbach, and R. J. Stevenson. 2010. An ecological assessment of Great Lakes tributaries in the Michigan Peninsulas. Journal of Great Lakes Research **36**:505-519.

Seelbach, P. W., M. J. Wiley, P. A. Soranno, and M. T. Bremigan. 2002. Aquatic conservation planning: using landscape maps to predict ecological reference conditoins for specific waters. Pages 454-478 *in* K. J. Gutzwiller, editor. Applying Landscape Ecologyin Biological Conservation. Springer-Verlag Publishers, New York.

Stevenson, R.J. 2006. Refining diatom indicators for valued ecological attributes and development of water quality criteria. In: Ognjanova-Rumenova, N. And K. Manoylov, eds. Advances in Phycological Studies. Pp. 365-383. Pensoft Publishers. Moscow.

Stevenson, R. J., R. C. Bailey, M. C. Harass, C. P. Hawkins, J. Alba-Tercedor, C. Couch, S. Dyer, F. A. Fulk, J. M. Harrington, C. T. Hunsaker, and R. K. Johnson. 2004. Designing data collection for ecological assessments. Pages 55-84 *in* M. T. Barbour, S. B. Norton, H. R. Preston, and K. W. Thornton, editors. Ecological Assessment of Aquatic Resources: Linking Science to Decision-Making. Society of Environmental Toxicology and Contamination Publication, Pensacola, Florida.

Stevenson, R. J., R. C. Bailey, M. C. Harass, C. P. Hawkins, J. Alba-Tercedor, C. Couch, S. Dyer, F. A. Fulk, J. M. Harrington, C. T. Hunsaker, and R. K. Johnson. 2004. Interpreting results of ecological assessments. Pages 85-111 *in* M. T. Barbour, S. B. Norton, H. R. Preston, and K. W. Thornton, editors. Ecological Assessment of Aquatic Resources: Linking Science to Decision-Making. Society of Environmental Toxicology and Contamination Publication, Pensacola, Florida.

Stevenson, R. J., B. E. Hill, A.T. Herlihy, L. L. Yuan, and S. B. Norton. 2008b. Algal-P relationships, thresholds, and frequency distributions guide nutrient criterion development. Journal of the North American Benthological Society **27**:783-799.

Stevenson, R. J., Y. Pan, K. Manoylov, C. Parker, D. P. Larsen, and A. T. Herlihy. 2008a. Development of diatom indicators of ecological conditions for streams of the western United States. Journal of the North American Benthological Society **27**:1000-1016.

Stevenson, R.J., J. Zalack, J. Wolin. submitted. A multimetric index of lake diatom condition using surface sediment assemblages. Freshwater Science

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

Table E-1. Candidate metric variable names, descriptions, and sources of traits. A check in the MMI column indicates the metric was included in the traditional and modeled MMIs. The traits source describes

MMI	Metrics Code	Description	Traits Source	Percentage basis
√	pT_distmmiSens	proportion of diatom Taxa characteristically found at minimally disturbed sites in the Western EMAP project	Stevenson et al (2008) WEMAP	Taxa with known characterizations from trait source
	pI_distmmiSens	proportion of diatom Individuals characteristically found at minimally disturbed sites in the Western EMAP project		
	pIsens_Achnanthaceae	proportion of Individuals in Taxa characteristically found at minimally disturbed sites in the Western EMAP project in the diatom family Achnanthaceae	Stevenson et al (2008) WEMAP	Taxa with known characterizations from trait source
	pIsens_Diatomaceae	proportion of Individuals in Taxa characteristically found at minimally disturbed sites in the Western EMAP project in the diatom family Diatomaceae	Stevenson et al (2008) WEMAP	Taxa with known characterizations from trait source
	pIsens_Naviculaceae	proportion of Individuals in Taxa characteristically found at minimally disturbed sites in the Western EMAP project in the diatom family Naviculaceae	Stevenson et al (2008) WEMAP	Taxa with known characterizations from trait source
	pItol_Naviculaceae	proportion of Individuals in Taxa characteristically found at highly disturbed sites in the Western EMAP project in the diatom family Naviculaceae	Stevenson et al (2008) WEMAP	Taxa with known characterizations from trait source
	pItol_Nitzschiaceae	proportion of Individuals in Taxa characteristically found at highly disturbed sites in the Western EMAP project in the diatom family Nitzschiaceae	Stevenson et al (2008) WEMAP	Taxa with known characterizations from trait source
√	pTsens_Achnanthaceae	proportion of Taxa characteristically found at minimally disturbed sites in the Western EMAP project in the diatom family Achnanthaceae	Stevenson et al (2008) WEMAP	Taxa with known characterizations from trait source
√	pTsens_Diatomaceae	proportion of Taxa characteristically found at minimally disturbed sites in the Western EMAP project in the diatom family Diatomaceae	Stevenson et al (2008) WEMAP	Taxa with known characterizations from trait source
√	pTsens_Naviculaceae	proportion of Taxa characteristically found at minimally disturbed sites in the Western EMAP project in the diatom family Naviculaceae	Stevenson et al (2008) WEMAP	Taxa with known characterizations from trait source
√	pTtol_Naviculaceae	proportion of Taxa characteristically found at highly disturbed sites in the Western EMAP project in the diatom family Naviculaceae	Stevenson et al (2008) WEMAP	Taxa with known characterizations from trait source
√	pTtol_Nitzschiaceae	proportion of Taxa characteristically found at highly disturbed sites in the Western EMAP project in the diatom family Nitzschiaceae	Stevenson et al (2008) WEMAP	Taxa with known characterizations from trait source
	pIPlnk_CelA	proportion of cells of all algae that are in Taxa characteristically found in plankton	Stevenson BPJ ecomorph traits file	All taxa in NRSA
√	pIHMotD_vDia	proportion of valves of diatoms that are in Taxa characterized as highly motile in sediments	Stevenson BPJ ecomorph traits file	All taxa in NRSA

MMI	Metrics Code	Description	Traits Source	Percentage basis
√	pIStlk_vDia	proportion of valves of diatoms that are in Taxa characterized as highly stalked	Stevenson BPJ ecomorph traits file	All taxa in NRSA
	pIFil_BioA	proportion of algal biovolume as filamentous taxa	Stevenson BPJ ecomorph traits file	All taxa in NRSA
	pICol_BioA	proportion of algal biovolume as colonial taxa	Stevenson BPJ ecomorph traits file	All taxa in NRSA
	pIPlnk_BioA	proportion of algal biovolume as planktonic taxa	Stevenson BPJ ecomorph traits file	All taxa in NRSA
	pIFil_CelA	proportion of algal cells as filamentous taxa	Stevenson BPJ ecomorph traits file	All taxa in NRSA
	pICol_CelA	proportion of algal cells as colonial taxa	Stevenson BPJ ecomorph traits file	All taxa in NRSA
	pIMotD_vDia	proportion of diatom valves as motile taxa	Stevenson BPJ ecomorph traits file	All taxa in NRSA
√	pIPlnk_vDia	proportion of diatom valves as planktonic taxa	Stevenson BPJ ecomorph traits file	All taxa in NRSA
	Achnanthidium	proportion of diatom valves in the genus *Achnanthidium*	Stevenson BPJ ecomorph traits file	All taxa in NRSA
	Cymbella	proportion of diatom valves in the genus *Cymbella*	Stevenson BPJ ecomorph traits file	All taxa in NRSA
	Cymbella_lato	proportion of diatom valves in the genus *Cymbella sensu lato*	Stevenson BPJ ecomorph traits file	All taxa in NRSA
	Nitzschia	proportion of diatom valves in the genus *Nitzschia*	Stevenson BPJ ecomorph traits file	All taxa in NRSA
	Planothidium	proportion of diatom valves in the genus *Planothidium*	Stevenson BPJ ecomorph traits file	All taxa in NRSA
	noNRSAT_units	number of NRSA taxa per natural unit counted		
	ShanDiv_units	Shannon diversity of non-diatom algal assemblages based on natural units counted		
	ShanDiv_cells	Shannon diversity of non-diatom algal assemblages based on cells counted		
	CountofOrder	number of non-diatom orders		

MMI	Metrics Code	Description	Traits Source	Percentage basis
	nOrder_units	number of non-diatom orders per natural unit counted		
	nFamily_units	number of non-diatom families per natural unit counted		
	nT_Diat	number of taxa of diatoms observed		
√	S_N_Diat	number of NRSA diatom taxa per valve counted		
√	ShanDiv_Diat	Shannon diversity of diatom assemblages		
	nOrders_valve	number of diatom orders per valve counted		
√	nFamilies_valve	number of diatom families per valve counted		
	nGenera_valve	number of diatom genera per valve counted		
	Plive	percent live diatoms		Live diatoms/(Live and dead diatoms)

Table E-2. Coefficients of determination (r2) of relationships between metrics and predictor variables of natural landscape features using random forest models (rf) and multiple linear regression (mlr), plus the t-statistics (ts) for comparisons of metrics in reference and highly disturbed sites when metrics were not modeled (ts.unm) and when metrics were modeled using random forest (ts.rf) and multiple linear regression (ts.mlr). Metrics for metric codes are explained in Table E-1. Coefficients of determination and t-statistics for metrics were calculated using one sample from all probability sites across the US (National) and from each of the three ecoregions (EHIGH = Eastern Highlands, PLNLOW = Lowland Plains, and WMTNS = Western Mountains).

Metrics	National					EHIGH					PLNLOW					WMTNS				
	r2.rf	r2.mlr	ts.unm	ts.rf	ts.mlr	r2.rf	r2.mlr	ts.unm	ts.rf	ts.mlr	r2.rf	r2.mlr	ts.unm	ts.rf	ts.mlr	r2.rf	r2.mlr	ts.unm	ts.rf	ts.mlr
pT_distmmiSens	0.52	0.52	11.51	10.57	1.01	0.08	0.53	10.89	8.64	1.00	0.32	0.39	9.42	7.39	2.20	0.61	0.81	7.61	6.67	-2.31
pI_distmmiSens	0.51	0.49	12.54	12.50	1.01	-0.21	0.16	13.88	11.84	14.07	0.37	0.39	10.21	8.95	2.13	0.45	0.66	7.68	5.88	0.53
pT_nutmmiSens	0.54	0.52	11.51	10.92	1.02	0.08	0.52	11.26	9.24	1.01	0.33	0.39	9.08	7.02	2.35	0.61	0.77	7.20	6.08	-1.86
pI_nutmmiSens	0.53	0.50	12.65	12.63	1.01	-0.29	0.25	15.31	12.92	1.01	0.28	0.40	9.89	8.18	1.83	0.43	0.65	7.88	5.86	0.63
pT_l1ptlSens	0.46	0.48	11.14	9.56	1.02	0.05	0.48	10.52	8.92	1.00	0.24	0.39	8.67	6.26	0.85	0.60	0.78	6.36	5.11	-2.68
pI_l1ptlSens	0.49	0.47	13.45	13.39	1.02	-0.20	0.30	16.61	14.38	1.01	0.21	0.32	10.10	8.22	-0.03	0.35	0.64	6.82	5.85	-0.30
pIFil_BioA	-0.10	0.08	-2.03	-1.88	1.84	0.05	0.45	0.21	0.34	-1.00	-0.15	0.15	0.84	1.14	2.48	-0.24	0.18	-1.65	-2.01	-1.17
pICol_BioA	-0.05	0.15	-0.83	-0.21	-0.54	-0.21	0.31	-1.10	-1.85	-1.54	0.05	0.21	-1.59	-0.39	-1.14	-0.09	0.54	0.31	1.12	-3.10
pIPlnk_BioA	-0.08	0.07	-1.82	-0.74	-2.01	-0.29	0.31	-2.15	-2.83	-2.37	-0.19	0.09	-1.48	0.95	-0.98	-0.20	0.51	-3.16	-1.87	-1.68
pIFil_CelA	0.01	0.09	-1.24	-0.36	-1.17	0.02	0.41	-0.39	-0.63	0.40	0.04	0.05	0.17	1.38	0.50	-0.27	0.24	-1.21	-1.55	-2.82
pICol_CelA	-0.08	0.20	-1.67	-0.19	2.50	-0.16	0.25	-2.56	-2.04	-2.27	-0.03	0.28	-0.79	0.48	2.82	-0.11	0.54	-0.36	0.86	-3.36
pIPlnk_CelA	-0.03	0.17	-6.00	-2.83	-0.99	-0.17	0.40	-4.33	-3.62	-1.00	-0.07	0.16	-4.19	-0.53	1.62	0.00	0.68	-4.05	-2.14	-1.26
pIMotD_vDia	0.07	0.22	1.06	-1.45	0.98	0.08	0.51	0.90	1.16	0.12	0.06	0.28	0.97	-1.32	-3.25	0.10	0.63	1.33	1.72	4.90
pIHMotD_vDia	0.38	0.38	-9.94	-9.20	-7.53	-0.16	0.41	-13.88	-13.19	-1.00	0.26	0.31	-7.91	-6.30	-5.79	-0.09	0.28	-5.28	-2.50	-1.97
pIHAdnt_vDia	0.23	0.31	1.06	-0.64	0.94	-0.14	0.58	-1.19	-0.08	-2.70	0.31	0.39	2.80	1.30	-0.36	-0.13	0.37	0.97	-0.46	-2.73
pIStlk_vDia	0.11	0.17	4.24	2.65	1.51	-0.08	0.35	1.35	1.87	1.00	0.00	0.25	3.17	1.67	1.50	0.00	0.45	3.00	2.04	-2.55
pIPlnk_vDia	-0.06	0.27	-5.96	-1.89	1.71	0.06	0.44	-4.57	-4.97	-2.83	0.09	0.37	-4.51	0.67	2.63	-0.17	0.56	-2.69	2.08	-3.45
Achnanthidium	0.36	0.48	8.72	7.99	1.05	-0.06	0.32	10.24	8.29	7.98	-0.12	0.28	6.55	4.75	1.78	0.09	0.52	2.86	2.71	4.54
Cymbella	-0.03	0.06	3.17	2.75	1.00	-0.03	0.88	1.65	1.44	1.00	-0.01	0.14	3.23	2.40	2.31	-0.20	0.31	0.68	1.10	0.59
Encyonema	0.07	0.17	4.22	2.31	1.44	-0.31	0.19	1.18	1.38	1.44	0.28	0.38	3.22	1.77	-0.16	-0.25	0.22	2.04	1.70	-1.64
Encyonopsis	-0.15	0.12	2.42	3.04	1.79	-0.31	0.16	2.22	1.64	5.25	-0.04	0.34	1.97	1.75	2.26	-0.16	0.24	-1.32	-0.23	1.32
Cymbella_lato	-0.03	0.09	5.30	4.07	3.81	-0.05	0.80	2.67	2.70	1.00	0.02	0.25	4.06	2.78	1.89	-0.25	0.31	1.72	1.97	-1.55
Nitzschia	0.29	0.32	-8.43	-6.95	-6.49	-0.06	0.43	-12.27	-11.72	-1.00	0.18	0.28	-7.51	-5.54	-5.22	-0.19	0.22	-4.89	-2.76	-1.76

Planothidium	0.03	0.13	0.34	-0.38	-1.01	0.05	0.28	-2.98	-1.91	-0.31	0.08	0.32	1.69	0.72	0.55	-0.27	0.41	1.06	-0.18	-2.74
pTsens_ Achnanthaceae	0.01	0.12	-7.60	-5.88	0.98	0.19	0.56	-4.14	-2.15	-2.74	0.03	0.24	-6.04	-4.53	-4.51	0.13	0.52	-1.75	-0.61	0.21
pTsens_ Diatomaceae	0.39	0.46	6.95	4.39	5.48	0.43	0.67	6.13	4.22	1.00	0.13	0.32	4.91	2.43	0.84	0.52	0.80	3.92	2.04	-3.78
pTsens_ Naviculaceae	0.22	0.29	11.28	8.94	4.65	-0.27	0.28	7.22	6.14	8.59	0.11	0.32	9.38	6.45	7.23	0.25	0.47	7.23	4.95	-1.29
pIsens_ Achnanthaceae	0.37	0.40	8.62	7.58	1.00	-0.10	0.28	9.58	7.67	5.80	0.17	0.29	7.10	6.54	1.85	0.04	0.48	4.63	2.70	2.97
pIsens_ Diatomaceae	0.13	0.23	3.92	2.60	3.48	0.20	0.56	3.70	2.27	4.33	0.11	0.34	2.54	1.50	-0.39	0.18	0.62	1.86	0.14	-3.76
pIsens_ Naviculaceae	0.10	0.19	7.20	5.81	4.68	-0.18	0.40	2.03	2.93	1.00	0.04	0.17	6.17	4.11	2.85	-0.14	0.38	3.39	3.54	0.25
pTtol_ Naviculaceae	0.37	0.37	-7.57	-6.35	-4.40	-0.16	0.44	-9.11	-7.77	-1.00	0.16	0.28	-3.18	-1.38	0.32	0.35	0.50	-6.39	-4.26	-1.41
pTtol_ Nitzschiaceae	0.26	0.36	-7.64	-4.94	-4.92	0.14	0.57	-7.15	-6.08	-1.00	0.22	0.39	-5.43	-3.20	-3.12	0.08	0.47	-4.08	-2.04	1.33
pItol_ Naviculaceae	0.37	0.29	-8.43	-7.35	-1.04	-0.04	0.33	-12.06	-11.58	-10.98	0.10	0.21	-4.65	-3.27	-2.18	0.41	0.52	-6.89	-5.71	-5.01
pItol_ Nitzschiaceae	0.28	0.30	-8.97	-7.19	-6.75	-0.12	0.49	-13.11	-12.53	-1.00	0.22	0.40	-6.82	-5.65	-5.02	-0.04	0.22	-4.44	-1.66	-0.98
noNRSAT_units	-0.02	0.11	-4.19	-2.72	-1.02	-0.18	0.56	-3.54	-2.53	-1.00	-0.17	0.29	-1.28	-0.37	-1.25	-0.14	0.23	-0.98	0.53	-0.40
noNRSAT_cells	-0.12	0.08	-3.72	-1.99	-3.14	-0.14	0.63	-1.86	-0.65	-1.00	-0.15	0.34	-0.58	0.37	-1.42	-0.08	0.33	-1.48	0.27	0.29
ShanDiv_units	0.18	0.24	-1.69	-0.48	-0.99	0.04	0.42	-5.64	-4.84	-4.05	-0.08	0.15	1.27	2.20	4.42	-0.10	0.45	-1.84	-0.47	-3.26
ShanDiv_cells	0.11	0.23	-2.24	-0.07	-0.99	0.14	0.54	-2.39	-1.62	-1.17	-0.05	0.16	0.14	0.57	3.43	-0.21	0.59	-3.18	-1.27	-1.93
CountOfOrder	0.08	0.21	0.54	1.36	-0.98	-0.03	0.41	-0.11	-0.11	-1.00	0.11	0.25	1.60	2.39	5.36	0.04	0.57	-0.31	-1.64	-3.42
nOrder_units	-0.02	0.19	-2.99	-1.52	-1.00	-0.02	0.60	-1.11	-0.20	-0.36	-0.07	0.34	-1.30	-0.48	0.53	-0.32	0.00	-0.02	0.69	-0.02
nFamily_units	-0.03	0.14	-3.40	-1.93	-3.20	-0.01	0.62	-1.81	-0.99	-0.92	-0.17	0.34	-1.35	-0.56	-0.24	-0.22	0.13	-0.03	0.88	0.41
nGenera_units	-0.03	0.10	-3.70	-2.32	-3.25	-0.17	0.60	-2.80	-1.78	-1.00	-0.20	0.27	-1.13	-0.24	-1.22	-0.16	0.26	-0.15	0.98	0.50
nT_Diat	0.20	0.14	-2.44	-1.42	-1.00	0.18	0.54	-4.53	-3.91	-1.00	0.11	0.27	1.23	3.00	4.18	-0.09	0.47	-2.52	-1.14	-0.22
S_N_Diat	0.07	0.10	-4.12	-2.60	-1.01	0.07	0.34	-4.11	-2.57	-4.40	0.05	0.27	-1.87	-0.71	-0.68	-0.09	0.48	-3.68	-2.62	-1.22
ShanDiv_Diat	0.21	0.25	-3.57	-2.12	-1.00	0.00	0.39	-6.52	-5.38	-1.00	-0.05	0.19	0.44	1.62	-1.53	-0.18	0.43	-2.46	-1.00	-3.18
nOrders_valve	-0.08	0.04	-1.45	-0.46	-0.84	-0.03	0.46	-0.72	0.94	-0.30	0.01	0.41	-2.33	-2.02	-1.36	-0.11	0.25	-3.19	-2.34	-2.50
nFamilies_valve	-0.03	0.04	-2.45	-1.42	-2.00	-0.12	0.42	-0.51	1.09	-1.16	0.07	0.30	-2.38	-2.24	-2.11	-0.01	0.23	-3.03	-2.33	-3.04
nGenera_valve	0.00	0.04	-2.74	-1.52	-1.01	-0.09	0.35	-2.09	-0.40	-2.38	0.01	0.27	-2.09	-1.72	-1.62	0.02	0.49	-2.10	-1.74	1.87

plive	0.42	0.39	-1.28	-1.95	-1.01	0.52	0.78	1.50	-2.64	0.89	0.30	0.41	-2.77	-3.49	-2.99	-0.02	0.28	-0.01	-0.49	-1.10

Table E-3. Linear model relationships between MMIs and percent watershed disturbed by human activities for different MMI models and for either all sites in the United States or one of three ecoregions (EHIGH = Eastern Highlands, PLNLOW = Lowland Plains, WMTNS = Western Mountains).

Model	Sites	Constant	Coefficient	F-stat	r2
Traditional	National	63.574	-5.286	568.3	0.217
Modeled	National	54.814	-2.393	149.2	0.068
Traditional	EHIGH	68.730	-6.868	137.7	0.202
Modeled	EHIGH	71.432	-6.547	130.7	0.194
Traditional	PLNLOW	58.877	-3.237	94.2	0.080
Modeled	PLNLOW	53.458	-1.314	19.3	0.017
Traditional	WMTNS	55.151	-5.312	46.6	0.100
Modeled	WMTNS	58.057	-4.007	33.4	0.073

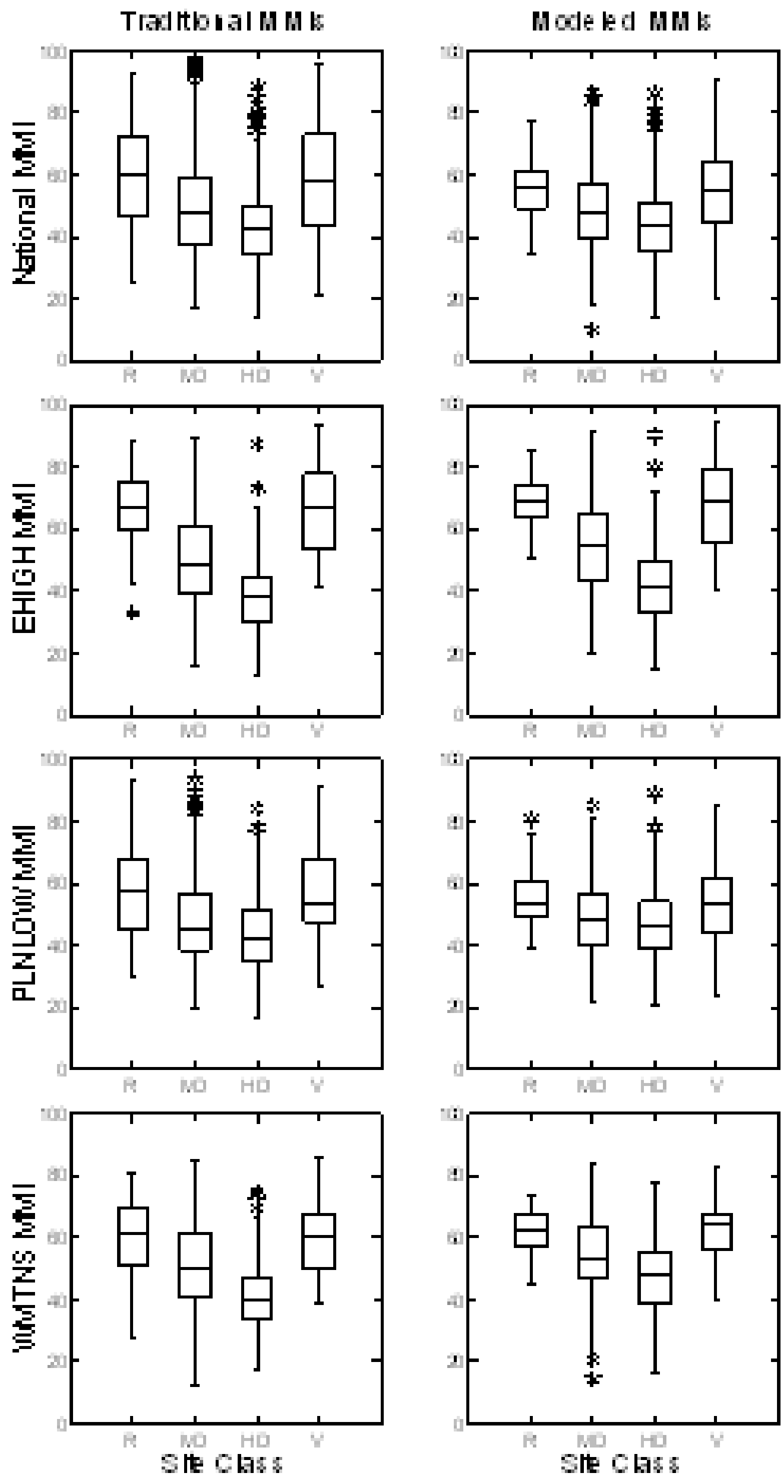


Figure E-1. Traditional and modeled MMIs for the national and ecoregion-specific scales compared among reference (R), moderately disturbed (MD), highly disturbed (HD), and validation (V) sites. EHIGH = Eastern Highland ecoregion; PLNLOW = Lowland Plains ecoregion; WMTNS = Western Mountains ecoregion.

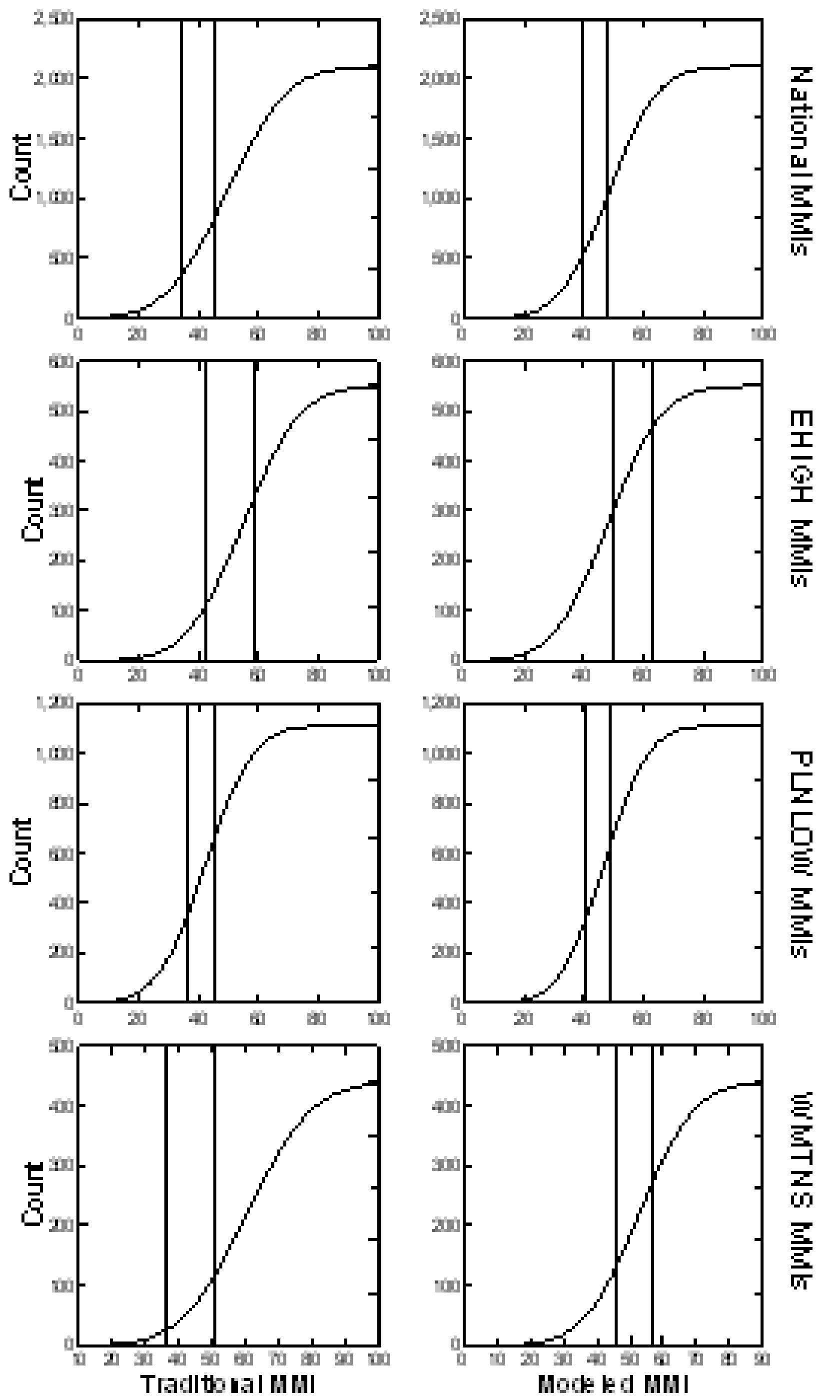


Figure E-2. Cumulative frequency distributions of traditional and modeled MMIs for the national and ecoregion-specific scales. The vertical dashed lines indicate, from left to right, the 5th and 25th percentiles of reference condition. Tic-marks of right side of the figures are the percentage of sites marked off in intervals of 20%. EHIGH = Eastern Highland ecoregion; PLNLOW = Lowland Plains ecoregion; WMTNS = Western Mountains ecoregion.

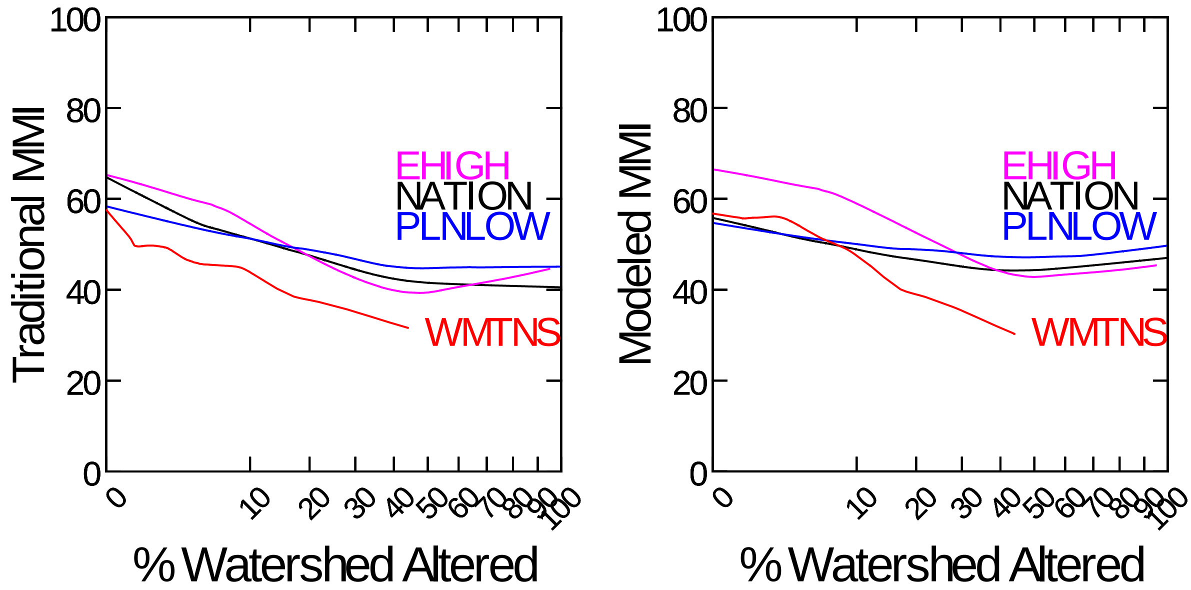


Figure E-3. The relationship between traditional and modeled MMIs related to percent watershed disturbed. EHIGH = Eastern Highland ecoregion; PLNLOW = Lowland Plains ecoregion; WMTNS = Western Mountains ecoregion.

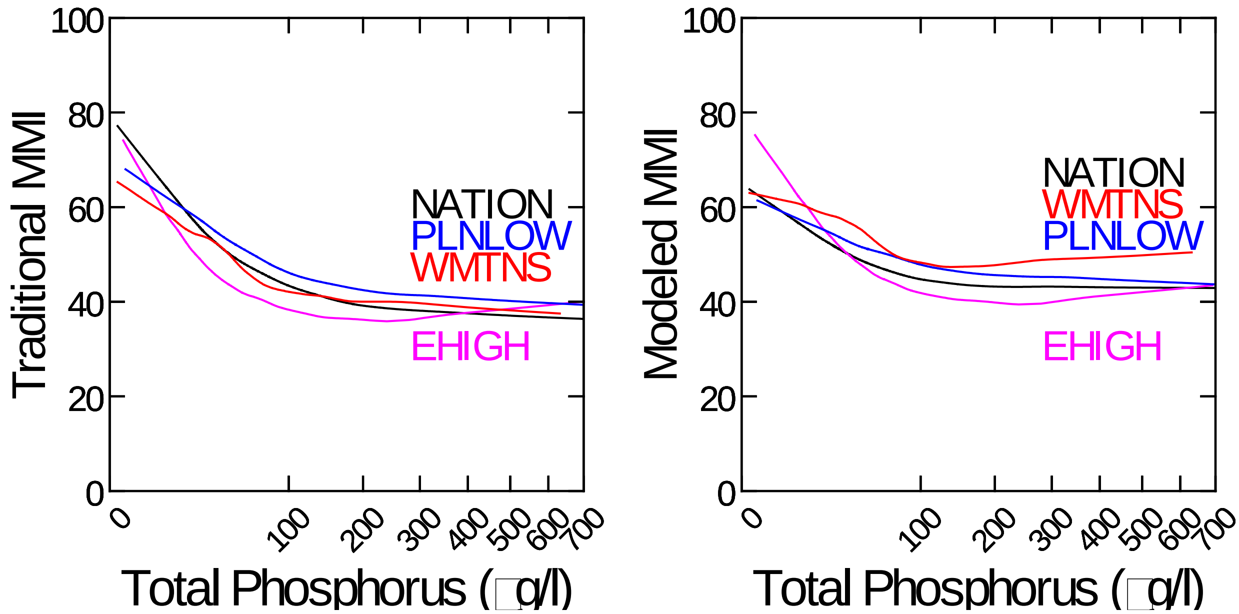


Figure E-4. The relationship between traditional and modeled MMIs related to total phosphorus concentration. EHIGH = Eastern Highland ecoregion; PLNLOW = Lowland Plains ecoregion; WMTNS = Western Mountains ecoregion.

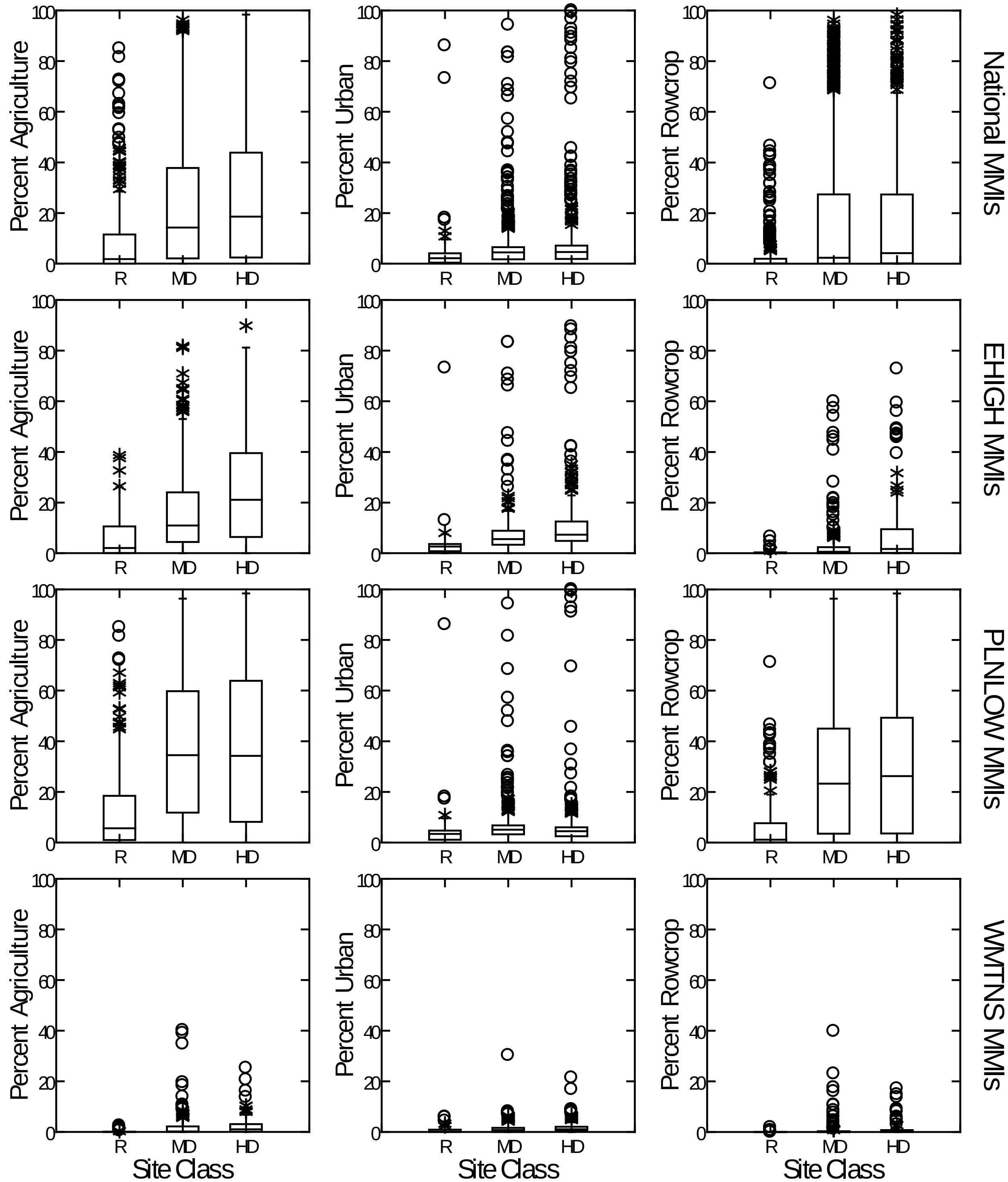


Figure E-5. Land use in watershed of reference (R), moderately disturbed (MD), and highly disturbed (HD) sites at the national scale and for each ecoregion.

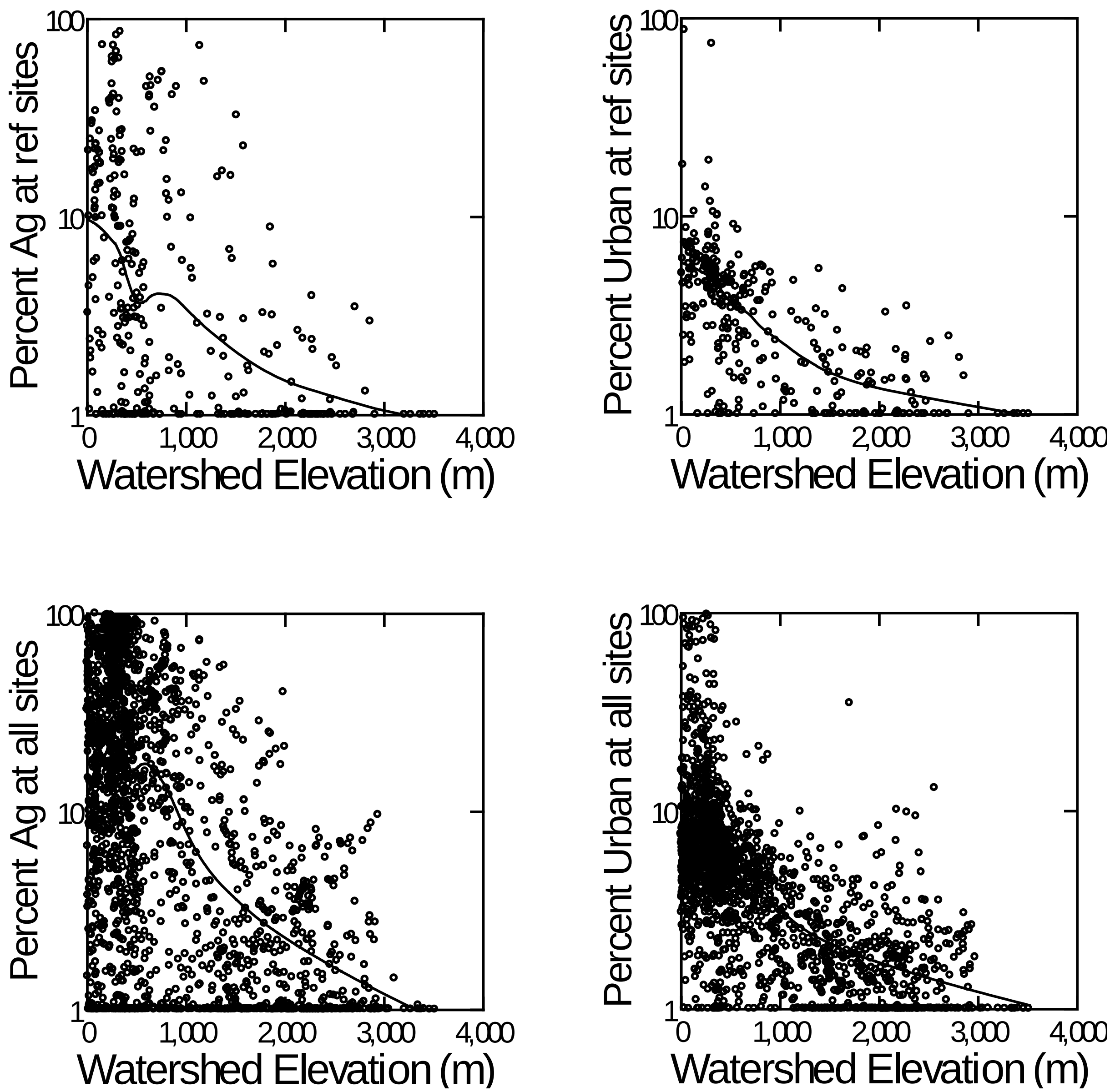


Figure E-6. Relationship between percentages of agricultural or urban land use within watersheds (with 1 added to enable log transformation of land use) for all sites and for reference sites only.

Water Chemistry Analysis for NRSA 2008–2009

6.1 Background information

The NRSA report summarizes four chemical stressors: total nitrogen, total phosphorus, acidity, and salinity. Criteria values and class definitions were identical to those used in the Wadeable Streams Assessment (WSA) as described below. Thresholds were established for each of the nine ecoregions reported on in the NRSA and WSA.

6.2 Threshold development

6.2.1 *Acidity and salinity*

For acidity, criteria values were determined based on values derived during the NAPAP program. Sites with acid neutralizing capacity (ANC) less than zero were considered acidic. Acidic sites with dissolved organic carbon (DOC) greater than 10 mg/L were classified as organically acidic (natural). Acidic sites with DOC less than 10 and sulfate less than 300 μeq/L were classified as acidic deposition impacted, while those with sulfate above 300 μeq/L were considered acid mine drainage impacted. Sites with ANC between 0 and 25 μeq/L and DOC less than 10 mg/L were considered acidic-deposition-influenced but not currently acidic. These low ANC sites typically become acidic during high flow events (episodic acidity).

Salinity and nutrient classes were divided into good, fair, or poor classes. Salinity classes were defined by specific conductance using ecoregional specific values (Table F-1).

6.2.2 *Total nitrogen and phosphorus*

Total nitrogen and phosphorus were classified using a method similar to that used for macroinvertebrate IBI classes using deviation from reference by aggregate ecoregion. For nutrients, the value at the 25th percentile of the reference distribution was selected for each region to define the least-disturbed condition class (good–fair boundary). The 5th percentile of the reference distribution defines the most disturbed condition class (Table F-1). For setting nutrient class boundaries, only reference sites from the screened WSA dataset were used. Since nutrients were the focus, the two nutrient criteria used in defining reference sites were dropped and the other seven criteria used by themselves to identify a set of “nutrient reference sites.” Before calculating percentiles from this set of sites, outliers (values outside 1.5 times the interquartile range) were removed. Percentiles were calculated in WSA from WSA data. They were not updated or changed based on NRSA data.

Table F-1. Nutrient and Salinity Category Criteria for NRSA Assessment

Ecoregion	Salinity as Conductivity (µS/cm) Good-Fair	Salinity as Conductivity (µS/cm) Fair-Poor	Total N (µg/L) Good-Fair	Total N (µg/L) Fair-Poor	Total P (µg/L) Good-Fair	Total P (µg/L) Fair-Poor
CPL	500	1000	1092	2078	56.3	108
NAP	500	1000	329	441	8.2	15.7
SAP	500	1000	296	535	17.8	24.4
UMW	500	1000	716	1300	21.6	44.7
TPL	1000	2000	1750	3210	165	338
NPL	1000	2000	948	1570	91.8	183
SPL	1000	2000	698	1570	52.0	95.0
WMT	500	1000	131	229	14.0	36.0
XER	500	1000	246	462	35.5	70.0

6.3 Literature cited

A.T. Herlihy, J.C. Sifneos. 2008. Developing nutrient criteria and classification schemes for wadeable streams in the conterminous US. Journal of the North American Benthological Society. Vol. 27, Issue 4.

Physical Habitat Assessment for the NRSA 2008–2009

7.1 Background information

An assessment of river and stream (fluvial) physical habitat condition is a major component of the NRSA. Of many possible general and specific fluvial habitat indicators measured in the NRSA surveys in 2008–2009, the assessment team chose streambed stability and excess fine sediments, instream habitat cover complexity, riparian vegetation, and riparian human disturbances for its assessment. These four indicators are generally important throughout the U.S. Furthermore, the project team had reasonable confidence in factoring out natural variability to determine expected values and the degree of anthropogenic alteration of the habitat attributes represented by these indicators.

In the broadest sense, fluvial habitat includes all physical, chemical, and biological attributes that influence or sustain organisms within streams or rivers. We use the term *physical habitat* to refer to the structural attributes of habitat. The NRSA made field measurements aimed at quantifying eight general attributes of physical habitat condition, including direct measures of human disturbance.

- Habitat volume/stream size
- Habitat complexity and cover for aquatic biota
- Streambed particle size
- Bed stability and hydraulic conditions
- Channel-riparian and floodplain interaction
- Hydrologic regime
- Riparian vegetation cover and structure
- Riparian disturbance

These attributes were previously identified during EPA's 1992 national stream monitoring workshop (Kaufmann 1993) as those essential for evaluating physical habitat in regional monitoring and assessments. They are typically incorporated in some fashion in regional habitat survey protocols (Platts et al. 1983, Fitzpatrick et al. 1998, Lazorchak et al. 1998, Peck et al. 2006, Peck et al. in press, EPA 2004) and were applied in the previous National Wadeable Streams Assessment (WSA) and the Western Rivers and Streams Pilot (EMAP-W) surveys conducted between 2000 and 2005 (EPA 2006, Stoddard et al. 2005a,b). The major habitat metrics used in those past assessments and considered in the NRSA are listed and defined in Table G-1. Some measures of these attributes are useful measures of habitat condition in their own right (e.g., channel incision as a measure of channel-riparian interaction); others are important controls on ecological processes and biota (bed substrate size), still others are important in the computation of more complex habitat condition metrics (e.g., bankfull depth is used to calculate relative bed stability [RBS]). Like biological characteristics, most habitat attributes vary according to their geomorphic and ecological setting. Even direct measures of riparian human activities and disturbances are strongly influenced by their geomorphic setting. And even within a region, differences in precipitation, stream drainage area channel gradient (slope) lead to variation in many aspects of stream habitat, because those factors influence discharge, flood stage, stream power (the product of discharge times gradient), and bed shear stress (proportional to the product of depth and slope). However, all eight of the major habitat attributes can be directly or indirectly altered by anthropogenic activities.

The NRSA follows the precedent of the EMAP-W and the WSA in reporting the condition of fluvial physical habitat condition on the basis of four habitat indicators that are important nationwide, can be reliably and economically measured, and their reference condition under minimal anthropogenic disturbance can be interpreted with reasonable confidence. These are: RBS as an indicator of bed sedimentation or hydrologic alteration, the areal cover and variety of fish concealment features as a measure of in-stream habitat complexity, riparian vegetation cover and structure as an indicator of riparian vegetation condition, and a proximity-weighted tally of streamside human activities as an indicator of riparian human disturbances (Paulsen et al., 2008).

In this document, we describe the approach taken by the NRSA for assessing physical habitat condition in rivers and streams based on the four abovementioned indicators. We also examine the rationale, importance, and measurement precision of each of these indicators, including the analytical approach for estimating reference conditions for each. Reference conditions for each indicator were interpreted as their expected value in sites having the least amount of anthropogenic disturbance within appropriately stratified regions. In most cases, we also refine the expected values as a function of geoclimatic controlling factors within regions. Finally, we examine patterns of association between physical habitat indicators and anthropogenic disturbance by contrasting habitat indicator values in least-, moderately, and most-disturbed sites nationally and within regions.

7.2 Methods

7.2.1 *Physical habitat sampling and data processing*

In the wadeable streams sampled in the NRSA, field crews took measurements while wading the length of each sample reach (Peck et al. 2006); in non-wadeable rivers, these measurements were made from boats (Peck et al., in press). Physical habitat data were collected from longitudinal profiles and from 11 cross-sectional transects and streamside riparian plots evenly spaced along each sampled stream reach (U.S. EPA 2007). The length of each sampling reach was defined proportional to the wetted channel width and measurements were placed systematically along that length to represent the entire reach. Sample reach lengths were 40 times the wetted channel-width (ChW) long in wadeable streams, with a minimum reach length of 150 m for channels less than 3.5 m wide. In non-wadeable rivers, reach lengths were also set to 40 ChW with a maximum length of 2,000 m. Thalweg depth measurements (in the deepest part of channel), habitat classification, and mid-channel substrate observations were made at tightly spaced intervals; whereas channel cross-sections and shoreline-riparian stations for measuring or observing substrate, fish cover (concealment features), large woody debris, bank characteristics and riparian vegetation structure were spaced further apart. Thalweg (maximum) depth was measured at points evenly spaced every 0.4 ChW along these reaches to give profiles consisting of 100 measurements (150 in streams <2.5m wide). The tightly spaced depth measures allow calculation of indices of channel structural complexity, objective classification of channel units such as pools, and quantification of residual pool depth, pool volume, and total stream volume. Channel slope and sinuosity on non-wadeable rivers were estimated from 1:24,000-scale digital topographic maps.

In wadeable streams, wetted width was measured and substrate size and embeddedness were evaluated using a modified Wolman pebble count of 105 particles spaced systematically along 21 equally spaced cross-sections, in which individual particles were classified visually into seven size-classes plus bedrock, hardpan and other (e.g., organic material). The numbers of pieces of large woody debris in the bankfull channel were tallied in 12 size classes (3 length by 4 width classes) along the entire length of sample reaches. Channel incision and the dimensions of the wetted and bankfull stream channel were measured at 11 equally spaced transects. Bank characteristics and areal cover of fish concealment features were visually assessed in 10 m long instream plots centered on transects, while riparian vegetation structure, presence of large (legacy) riparian trees, non-native (alien) riparian plants, and evidence of human disturbances (presence/absence and proximity) in 11 categories were visually assessed on adjacent 10 × 10 m riparian plots on both banks. In addition, channel gradient (slope) in wadeable streams was measured to provide information necessary for calculating reach gradient residual pool volume and RBS. In wadeable streams, crews used laser or hydrostatic levels for slopes <2.5%, and optionally were allowed to use hand-held clinometers in channels with slopes >2.5%. Compass bearings between stations were obtained for calculating channel sinuosity. Channel constraint and evidence of debris torrents and major floods were assessed over the whole reach after the other components were completed. Discharge was measured by the velocity-area method at the time of sampling, or by other approximations if that method was not practicable (Peck et al. 2006; EPA 2007). Two-person crews typically completed NRSA habitat measurements in 1.5 to 4 hours of field time, though large, deep streams that were only marginally wadeable took up to several hours longer.

In non-wadeable rivers, NRSA field crews floating downstream in inflatable rafts, or in slower rivers small power boats, measured the longitudinal thalweg depth profile (approximated at mid-channel) using 7.5m telescoping survey rods or SONAR, at the same time tallying snags and off-channel habitats, classifying main channel habitat types, and characterizing mid-channel substrate by probing the bottom. At 11 littoral/riparian plots (each 10m wide x 20m long) spaced systematically and alternating sides along the river sample reach, field crews measured channel wetted width, bankfull channel dimensions, incision, channel constraint. They assessed near-shore, shoreline, and riparian physical habitat characteristics by measuring or observing littoral depths, riparian canopy cover, substrate, large woody debris, fish cover, bank characteristics, riparian vegetation structure, presence of large ("legacy") riparian trees, non-native riparian plants, and evidence of human activities. After all the thalweg and littoral/riparian measurements and observations were completed, the crews estimated the extent and type of channel constraint (see Peck et al. in press; EPA 2007). Channel slope and sinuosity on non-wadeable rivers were estimated from 1:24,000-scale digital topographic maps.

See Kaufmann et al. (1999) for calculations of reach-scale summary metrics from field data, including mean channel dimensions, residual pool depth, bed particle size distribution, wood volume, riparian vegetation cover and complexity, and proximity-weighted indices of riparian human disturbances. See Faustini and Kaufmann (2007) for details on the calculation of geometric mean streambed particle diameter, and Kaufmann et al. (2008, 2009) for calculations of bed shear stress and RBS that have been modified since published by Kaufmann et al. (1999), and Kaufmann and Faustini (2012) demonstrating the utility of EMAP and NRSA channel morphologic data to estimate transient storage and hydraulic retention in wadeable streams.

7.2.2 *Quantifying the precision of physical habitat indicators*

The absolute and relative precision of the physical habitat condition metrics used in the NRSA are shown in Table G-2, determined for most of the variables based on 2,113 unique sites and repeat visits to a random subset of 197 of those sites. The RMS_{rep} expresses the precision or replicability of field measurements, quantifying the average variation in a measured value between same-season site revisits, pooled across all sites where measurements were repeated. We calculated RMS_{rep} as the root-mean-square error of repeat visits during the same year, equivalent to the root mean-square error (RMSE) relative to the site means, as discussed Kaufmann et al. (1999) and Stoddard et al. (2005a). S/N is the ratio of variance among streams to that for repeat visits to the same stream as, described by Kaufmann et al. (1999).

The ability of a monitoring program to detect trends is sensitive to the spatial and temporal variation in the target indicators as well as the design choices for the network of sites and the timing and frequency of sampling. Sufficient temporal sampling of sites was not available to estimate all relevant components of variance for the entire U.S. However, Larsen et al. (2004) examined the components of sampling variability for a number of the EMAP physical habitat variables including some of interest in this paper (residual depth, canopy cover, fine sediment, and large wood). Their analysis was based on evaluation on six Pacific Northwest surveys that included 392 stream reaches and 200 repeat visits. These surveys were conducted in Oregon and Washington from 1993 to 1999. Most were from one to three years in duration, but one survey lasted six years. They modeled the likelihood of detecting a 1–2% per year trend in the selected physical habitat characteristics, if such a trend occurs, as a function of the duration of a survey. To calculate the number of years required to detect the defined trends in a monitoring network with a set number of sites, they set the detection probability at >80% with <5% probability of incorrectly asserting a trend if one is not present. We used the same survey data sets to duplicate their analysis for several variables not included in the Larsen et al. (2004) publication, including log transformed relative bed stability (LRBS_BW5) and riparian vegetation cover complexity (XCMGW, the combined cover of three layers of riparian woody vegetation); the results of that trend detection potential is summarized in Table G-3.

7.3 Physical habitat condition indicators in the NRSA

7.3.1 *Relative bed stability and excess fines*

Streambed characteristics (e.g., bedrock, cobbles, silt) are often cited as major controls on the species composition of macroinvertebrate, periphyton, and fish assemblages in streams (e.g., Hynes 1970, Cummins 1974, Platts et al. 1983, Barbour et al. 1999, Bryce et al., 2008, 2010). Along with bedform (e.g., riffles and pools), streambed particle size influences the hydraulic roughness and consequently the range of water velocities in a stream channel. It also influences the size range of interstices that provide living space and cover for macroinvertebrates and smaller vertebrates. Accumulations of fine substrate particles (excess fine sediments) fill the interstices of coarser bed materials, reducing habitat space and its availability for benthic fish and macroinvertebrates (Hawkins et al. 1983, Platts et al. 1983, Rinne 1988). In addition, these fine particles impede circulation of oxygenated water into hyporheic habitats reducing egg-to-emergence survival and growth of juvenile salmonids (Suttle et al. 2004). Streambed characteristics are often sensitive indicators of the effects of human activities on streams (MacDonald et al. 1991, Barbour et al. 1999, Kaufmann et al. 2009). Decreases in the mean particle size and increases in streambed fine sediments can destabilize stream channels (Wilcock 1997, 1998) and may indicate increases in the rates of upland erosion and sediment supply (Lisle 1982, Dietrich et al. 1989).

"Unscaled" measures of surficial streambed particle size, such as percent fines or D_{50}, can be useful descriptors of stream bed conditions. In a given stream, increases in percent fines or decreases in D_{50} may result from anthropogenic increases in bank and hillslope erosion. However, a great deal of the variation in bed particle size among streams is natural: the result of differences in stream or river size, slope, and basin lithology. The power of streams to transport progressively larger sediment particles increases in direct proportion to the product of flow depth and slope. All else being equal, steep streams tend to have coarser beds than similar size streams on gentle slopes. Similarly, the larger of two streams flowing at the same slope will tend to have coarser bed material, because its deeper flow has more power to scour and transport fine particles downstream (Leopold et al. 1964, Morisawa 1968). For these reasons, we "scale" bed particle size metrics, expressing bed particle size in each stream as a deviation from that expected as a result of its size, power, and landscape setting (Kaufmann et al., 1999, 2008, 2009).

The scaled median streambed particle size is expressed as RBS, calculated as the ratio of the geometric mean diameter, D_g, divided by D_{cbf}, the critical diameter (maximum mobile diameter) at bankfull flow (Gordon et al., 1992), where D_g is based on systematic streambed particle sampling ("pebble counts") and D_{cbf} is based on the estimated streambed shear stress calculated from slope, channel dimensions, and hydraulic roughness during bankfull flow conditions.

RBS is a measure of habitat stability for aquatic organisms as well as an indication of the potential for economic risk to streamside property and structures from stream channel movement. In many regions of the U.S., we may also be able to use RBS to infer whether sediment supply is augmented by upslope or bank erosion from anthropogenic or other disturbances, because it can indicate the degree of departure from a balance between sediment supply and transport. In interpreting RBS on a regional scale, Kaufmann et al. (1999, 2009) argued that, over time, streams and rivers adjust sediment transport to match supply from natural weathering and delivery mechanisms driven by the natural disturbance regime, so that RBS in appropriately stratified regional reference sites should tend towards a range characteristic of the climate, lithology, and natural disturbance regime. Values of the RBS index either substantially lower (finer, more unstable streambeds) or higher (coarser, more stable streambeds) than those expected based on the range found in least-disturbed reference sites within an ecoregion are considered to be indicators of ecological stress.

Excess fine sediments can destabilize streambeds when the supply of sediments from the landscape exceeds the ability of the stream to move them downstream. This imbalance results from numerous human uses of the landscape, including agriculture, road building, construction, and grazing. Lower-than-expected streambed stability may result either from high inputs of fine sediments (from erosion) or increases in flood magnitude or frequency (hydrologic alteration). When low RBS results from fine sediment inputs, stressful ecological conditions result from fine sediments filling in the habitat spaces between stream cobbles and boulders (Bryce et al. 2008, 2010). Instability (low RBS) resulting from hydrologic alteration can be a precursor to channel incision and arroyo formation (Kaufmann et al. 2009). Perhaps less well recognized, streams that have higher than expected streambed stability can also be considered stressed—very high bed stability is typified by hard, armored streambeds, such as those often found below dams where fine sediment flows are interrupted, or within channels where banks are highly altered. Values of RBS higher than reference expectations can indicate anthropogenic coarsening or armoring of streambeds, but streams containing substantial amounts of bedrock may also have very high RBS, and at this time it is difficult to determine the role of human alteration in stream coarsening on a national scale. For this reason, NRSA reported only on the "low end" of RBS relative to reference conditions, generally indicating stream bed excess fine sediments or augmented stormflows associated with human disturbance of stream drainages and riparian zones.

7.3.1.1 Precision of sediment and bed stability measurements

The geometric mean bed particle diameter (D_{gm}) and *RBS* varied over 8 orders of magnitude in the NRSA surveys. Because of this wide variation and the fact that both exhibit repeat-visit variation that is proportional to their magnitude at individual streams, it is useful and necessary to log transform these variables (LSUB_DMM and LRBS_g08). The RMS_{rep} of LSUB_DMM in wadeable streams of the EMAP-W survey was 0.246, similar to that reported by Faustini and Kaufmann (2007) for EMAP-W (0.21). For a D_{gm} = *"y"* mm, the log-based RMS_{rep} of 0.246 translates to an asymmetrical 1SD error bound of 0.57*y* to 1.76*y* mm. The RMS_{rep} of *LRBS_g08* in NRSA wadeable streams was 0.48, approximately 6% of its observed range, but less precise (surprisingly) than that for EMAP-W (RMS_{rep} = 0.365). The log-based RMS_{rep} of 0.48 for NRSA LRBS_g08 translates to an asymmetrical error bound of 0.33*y* to 3.0*y* around an untransformed RBS value of *"y"* (Table G-2). Compared with the high S/N ratio for LSUB_DMM in the NRSA

(12.4 for wadeable+boatable waters), relative precision for LRBS_g08 was lower (S/N=5.0), reflecting the reduction in total variance when a large component of natural variability is "modeled out" by scaling for channel gradient, water depth, and channel roughness. Nevertheless, the relative precision of LRBS_g08 is moderately high and easily adequate to make it a useful variable in regional and national assessments (Kaufmann et al. 1999, 2008, Faustini and Kaufmann 2007). The transformation of the unscaled geometric mean bed particle diameter D_{gm} to the ratio RBS by dividing by the critical diameter reduced the within-region variation by accounting for some natural controlling factors. As a result, we feel that the scaled variable helps to reveal alteration of bed particle size and mobility from anthropogenic erosion and sedimentation (Kaufmann et al. 2008, 2009).

We have examined the components of variability of LRBS based on earlier surveys and modeled its potential utility in trend detection in the Pacific Northwest region of the U.S. with the same data and procedures as used by Larsen et al. (2004), in which all methods were the same as used in EMAP-W and WSA except that bed substrate mean diameter data used by Larsen et al. was determined based on 55, rather than 105 particles. (NRSA data differed from data used in that analysis by using laser levels rather than hand-held clinometers to measure wadeable stream slopes <2/5%) That analysis showed that a 50-site monitoring program could detect a subtle trend in LRBS_BW5 of 2% per year within 8 years, if sites were visited every year (Table G-3).

7.3.2 *Instream habitat cover complexity*

Although the precise mechanisms are not completely understood, the most diverse fish and macroinvertebrate assemblages are usually found in streams that have complex mixtures of habitat features: large wood, boulders, undercut banks, tree roots, etc. (see Kovalenko et al. 2011). When other needs are met, complex habitat with abundant cover should generally support greater biodiversity than simple habitats that lack cover (Gorman and Karr 1978, Benson and Magnuson 1992). Human use of streams and riparian areas often results in the simplification of this habitat, with potential effects on biotic integrity (Kovalenko et al. 2011). For this assessment, we use a measure (XFC_NAT in Kaufmann et al., 1999) that sums the amount of instream habitat consisting of undercut banks, boulders, large pieces of wood, brush, and cover from overhanging vegetation within a meter of the water surface, all of which were estimated visually by NRSA field crews.

7.3.2.1 Quantifying habitat complexity

Habitat complexity is difficult to quantify, and could be quantified or approximated by a wide variety of measures. The NRSA Physical Habitat protocols provide estimates for nearly all of the following components of complexity identified during EPA's 1992 stream monitoring workshop (Kaufmann 1993):

- Habitat type and distribution (e.g., Bisson et al. 1982, O'Neill and Abrahams 1984, Frissell et al. 1986, Hankin and Reeves 1988, Hawkins et al. 1993, Montgomery and Buffington 1993, 1997, 1998).

- Large wood count and size (e.g., Harmon et al. 1986, Robison and Beschta 1989, Peck et al. 2006).
- In-channel cover: percentage areal cover of fish concealment features, including undercut banks, overhanging vegetation, large wood, boulders (Hankin and Reeves 1988, Kaufmann and Whittier 1997, Peck et al. 2006)
- Residual pools, channel complexity, hydraulic roughness (e.g., Kaufmann 1987a, b, Lisle 1987, Stack and Beschta 1989; Lisle and Hilton 1992, Robison and Kaufmann 1994, Kaufmann et al. 1999, 2008, Kiem et al. 2002, Kaufmann et al. 2011)
- Width and depth variance, bank sinuosity (Kaufmann 1987a, Moore and Gregory 1988, Kaufmann et al. 1999, Madej 1999, 2001, Kaufmann et al. 2008, Mossop and Bradford 2006, Pearsons and Temple, 2007, 2010, Kaufmann and Faustini, 2011).

Residual depth is a measure of habitat volume, but also serves as one of the indicators of channel habitat complexity, particularly when expressed as a deviation from reference expectations, including the influences of basin size. A stream with more complex bottom profile will have greater residual depth than one of similar drainage area, discharge, and slope that lacks that complexity (Kaufmann 1987a). Conversely, between two streams of equal discharge and slope, the one with greater residual depth (i.e., larger, more abundant residual pools) will have greater variation in cross-sectional area, slope, and substrate size. A related measure of the complexity of channel morphology is the coefficient of variation in thalweg depth, calculated entirely from the thalweg depth profile (SDDEPTH ÷ XDEPTH). The thalweg profile is a systematic survey of depth in the stream channel along the path of maximum depth ("thalweg"). In addition to measures of channel morphometric complexity, EMAP habitat protocols measure in-channel large wood (sometimes called "large woody debris" or simply "LWD"), and several estimates of the areal cover of various types of fish and macroinvertebrate "cover" or concealment features. The large wood metrics include counts of wood pieces per 100 m of bankfull channel and estimates of large wood volume in the sample reach expressed in cubic meters of wood per square meter of bankfull channel. The "fish cover" variables are visual estimates of the areal cover of single or combined types of habitat features.

NRSA required a general summary metric as a holistic indicator of many aspects of habitat complexity, so used the metric XFC_NAT, summing the areal cover from large wood, brush, overhanging vegetation, live trees and roots, boulders, rock ledges, and undercut banks in the wetted stream channel. Habitat complexity and the abundance of particular types of habitat features differ naturally with stream size, slope, lithology, flow regime, and potential natural vegetation. For example, boulder cover will not occur naturally in streams draining deep deposits of loess or alluvium that do not contain large rocks. Similarly, large wood will not be found naturally in streams located in regions where riparian or upland trees do not grow naturally. Though the combined cover index XFC_NAT partially overcomes these differences, we set stream-specific expectations for habitat complexity metrics in the NRSA based on region-specific reference sites and further refined them as a function of geoclimatic controls.

7.3.2.2 Precision of habitat complexity measures

The instream habitat complexity index XFC_NAT ranged from 0 to 2.3, or 0% to 230% in NRSA, expressing the combined areal cover of the five cover elements contributing to its sum. The RMS_{rep} of Log(0.01+XFC_NAT) was 0.24, meaning that an XFC_NAT value of 10% cover at a single stream site has a $\pm$1.0 RMS_{rep} error bound of 6% to 17% (Table G-2) S/N was relatively low for this indicator (1.87), though higher in wadeable streams (2.29) than in boatable rivers (1.22). Despite its relatively low S/N, the RMS_{rep} for LXFC_LWD was 10% of the observed range of XFC_NAT. It was retained as a habitat complexity indicator because it contains biologically relevant information not available in other metrics, showed moderate responsiveness to human disturbances, and has precision adequate to discern relatively large differences in habitat complexity.

7.3.3 ***Riparian vegetation***

7.3.3.1 Quantifying riparian vegetation cover complexity

The importance of riparian vegetation to channel structure, cover, shading, inputs of nutrients and large wood, and as a wildlife corridor and buffer against anthropogenic disturbance is well recognized (Naiman et al. 1988, Gregory et al. 1991). Riparian vegetation not only moderates stream temperatures through shading, but also increases bank stability and the potential for inputs of coarse and fine particulate organic material. Organic inputs from riparian vegetation become food for stream organisms and provide structure that creates and maintains complex channel habitat.

The presence of a complex, multi-layered vegetation corridor along streams and rivers is an indicator of how well the stream network is buffered against sources of stress in the watershed. Intact riparian areas can help reduce nutrient and sediment runoff from the surrounding landscape, prevent streambank erosion, provide shade to reduce water temperature, and provide leaf litter and large wood that serve as food and habitat for stream organisms (Gregory et al. 1991). The presence of large, mature canopy trees in the riparian corridor reflects its longevity, whereas the presence of smaller woody vegetation typically indicates that riparian vegetation is reproducing, and suggests the potential for future sustainability of the riparian corridor.

NRSA evaluated the cover and complexity of riparian vegetation based on the metric XCMGW, which is calculated from visual estimates made by field crews of the areal cover and type of vegetation in three layers: the ground layer (<0.5m), med-layer (0.5-5.0 m) and upper layer (>5.0 m). The separate measures of large and small diameter trees, woody and non-woody mid-layer vegetation, and woody and non-woody ground cover are all visual estimates of areal cover. XCMGW sums the cover of *woody* vegetation summed over these three vegetation layers, expressing both the abundance of vegetation cover and its structural complexity. Its theoretical maximum is 3.0 if there is 100% cover in each of the three vegetation layers. XCMGW gives an indication of the longevity and sustainability of perennial vegetation in the riparian corridor (Kaufmann et al. 1999).

7.3.3.2 Precision of riparian vegetation index

XCMGW ranged from 0 to 2.8 (280% cover), with RMS_{rep} of Log(0.01+XCMGW) = 0.146 (Table G-2), meaning that an XCMGW value of 10% at a single stream site has a $\pm$1.0 RMS_{rep} error bound of 7% to 14%. Its S/N ratio was 9.38, indicating very good potential for discerning differences among sites. We examined the components of variability of XCMGW and modeled its potential utility in trend detection in the Pacific Northwest region of the U.S. with the same data and procedures as used by Larsen et al. (2004). Based on that analysis, a 50-site monitoring program could detect a subtle trend in XCMGW of 2% per year within 8 years, if sites were visited every year (Table G-3).

7.3.4 ***Riparian human disturbances***

Agriculture, roads, buildings, and other evidence of human activities in and near the stream and river channel may exert stress on aquatic ecosystems and may also serve as indicators of overall anthropogenic stress. EPA's 1992 stream monitoring workshop recommended field assessment of the frequency and extent of both in-channel and near-channel human activities and disturbances (Kaufmann 1993). The vulnerability of the stream network to potentially detrimental human activities increases with the proximity of those activities to the streams themselves. The NRSA follows Stoddard et al. (2005b) and U.S. EPA (2006) in using a direct measure of riparian human disturbance that tallies 11 specific forms of human activities and disturbances (walls, dikes, revetments or dams; buildings; pavement or cleared lots; roads or railroads; influent or effluent pipes; landfills or trash; parks or lawns; row crop agriculture; pasture or rangeland; logging; and mining) at 22 separate locations along the stream reach, and weights them according to how close to the channel they are observed (W1_HALL in Kaufmann et al. 1999). Observations within the stream or on its banks are weighted by 1.5, those within the 10 $\times$ 10 meter plots are weighted by 1.0, and those visible beyond the plots are weighted by 0.5. The index W1_HALL ranged from 0 (no observed disturbance) to ~7 (e.g., equivalent to four or five types of disturbance observed in the stream, throughout the reach; or seven types observed within all 22 riparian plots bounding the stream reach). Although direct human activities certainly affect riparian vegetation complexity and layering measured by the riparian vegetation index (previous paragraph), the riparian disturbance index is more encompassing, and differs by being a *direct* measure of observable human activities that are presently or potentially detrimental to streams.

7.3.4.1 Precision of riparian disturbance indicators

W1_HALL ranged from 0 to 7.3 in the NRSA. The precision of the weighted human disturbance indicator W1_HALL was proportional to the level of disturbance. The RMS_{rep} of log(0.1+W1_HALL) was 0.186 (Table G-2), meaning that a W1_HALL value of 1.0 at a single stream site has a $\pm$1.0 RMS_{rep} error bound of 0.65 to 1.53. The relative precision of Log(0.1+W1_HALL) was moderate (S/N=5.18)

7.4 Estimating reference condition for physical habitat

7.4.1 *Reference site screening and anthropogenic disturbance classifications*

As part of the routine application of its field and GIS protocols, the NRSA obtained various measures of human disturbance associated with each site and its catchment. Site scale indicators of human disturbance included evidence of various human activities including nearby roads, riprap, agricultural activities, riparian vegetation disturbance, etc., as detailed by Kaufmann et al. (1999). These indicators of local scale disturbance were used in combination with water chemistry (chloride, total phosphorus, total nitrogen, and turbidity), as described by Herlihy et al. (2008), to screen probability and hand-picked sites and designate them as least-, moderately, and highly disturbed, relative to other sites within each of the regions of the NRSA. In addition, we used basin and sub-basin row crop and urban land use percentages and the density of dams and impoundments as described in the reference technical section to rank sites by disturbance categories, as shown in Table G-4. To avoid circularity, we did not use any field measures of sediment, in-channel habitat complexity, or riparian vegetation to screen least-disturbed sites used to estimate reference condition for excess streambed fining, instream fish cover, and riparian vegetation. Nor did we use such measures in defining levels of disturbance to use in examining the associations of these habitat metrics with human disturbances. We did, however, use field observations of the level and proximity of streamside human activities in screening reference sites and defining levels of disturbance for evaluating indicator responsiveness. In the table, the designation "**R**" refers to least-disturbed ("reference") sites; "**M**" to moderately disturbed sites, and "**D**" to the most-disturbed sites within each of the nine aggregate ecoregions discussed herein. We defined these site disturbance categories independent of the habitat indicators we evaluated (other than riparian human disturbances), allowing an assessment of fluvial habitat response to a gradient of human activities and disturbances.

7.4.2 *Modeling expected reference values of the indicators*

For LRBS, we modeled expected values based on the distribution of LRBS in reference sites within regions or groups of regions. In some regions boatable and wadeable rivers and streams are modeled separately; in others they are combined. Where possible, we used regression models in which W1_HALL (human disturbance) is set to zero in regressions of LRBS=f(W1_HALL) within reference sites only (RMD_PHAB=R). In these cases, the adjusted mean of the reference distribution is defined as the y-intercept of these regressions and the SD about the adjusted reference mean is defined as the RMSE of those regressions. Condition classes were defined based on normal approximation of the 5th and 25th percentiles of the actual or adjusted reference distributions. The definition of "poor" condition was set as those sites with LRBS below the reference mean LRBS minus 1.65(SD_{ref}). Sites in "good" condition with respect to this indicator were those with LRBS above the reference mean LRBS minus 0.67(SD_{ref}).

For instream fish cover complexity, we estimated expected XFC_NAT based on multiple linear regression models predicting Log_{10}(XFC_NAT) in reference sites from geoclimatic controlling factors within regions or aggregated regions. Because there is a gradient of human disturbance within the set of reference sites in all the regions considered, and it was correlated with XFC_NAT, we also incorporated field measures of human disturbance into the regressions. Site-specific expected ("E") values of XFC_NAT were then calculated by setting human the disturbance metric values to very low values (but never lower than observed among the reference). We then calculated observed/expected (O/E) values of XFC_NAT and examined their distribution among reference sites. Because we had modeled-out disturbance to some extent in our calculation of E values, the distribution of O/E in reference sites did not necessarily have a mean of 1/1 (Log=0), although means were very close to 1/1. We set expectations of the O/E values based on the mean and SD of the regional reference distributions, analogous to that described for LRBS in the previous paragraph.

For riparian condition—XCMGW transformed as L_xcmgw= Log_{10}(0.01+XCMGW)—we estimated expected condition based on simple regional reference site distributions or regression models in which W1_HALL was set to zero in regressions prediction L_xcmgw as a function of W1_HALL within the subset of reference sites (RMD_PHAB = R). The adjusted mean L_xcmgw for reference sites was defined as the y-intercept and the SD about the reference mean is defined as the RMSE of those regressions.

We did not base thresholds of riparian disturbance on the reference distributions, as was done for sediment, habitat complexity, and riparian vegetation condition. Rather, the classes for riparian disturbance were set using the same judgment-based criteria for all regions. W1_HALL, the database variable name for this indicator, is a direct measure of human disturbance "pressure"—unlike the other habitat indicators, which are actually measures of habitat response to human disturbance pressures. It is very difficult to define reference sites without screening sites based on these human disturbance tallies (i.e., W1_HALL). For this reason, we took a different approach for setting riparian disturbance thresholds, defining low disturbance sites as those with W1_HALL <0.33 and high riparian disturbance sites as those with W1_HALL >1.5; we applied these same thresholds in all ecoregions. A value of 1.5 for a stream means, for example, that at 22 locations along the stream the field crews found an average of one of 11 types of human disturbance within the stream or its immediate banks. A value of 0.33 means that, on average, one type of human disturbance was observed at one-third of the 22 riparian plots along a sample stream or river.

7.5 Response of the physical habitat indicators to human disturbance

The sedimentation and riparian vegetation indicators, LRBS and XCMGW, showed modest to strong negative response to human disturbance in most regions and aggregations of regions, as illustrated by t-values (+2.11-12.24) comparing differences in means of reference minus disturbed sites (Table G-5). However, sediment and riparian vegetation associations with human disturbance tended to be slightly stronger for sediments and much stronger for riparian vegetation in wadeable versus boatable sites (Table G-5, Figures G-1 and G-2).

Except for the weak contrary response in the Eastern Highlands (t = -1.26), the instream habitat complexity indicator showed moderate to moderate response to human disturbance, with t-values ranging from +2.13 to +4.25 (Table G-5). As for the other habitat indicators, associations were in most cases stronger for wadeable versus boatable sites.

Because the field-obtained measures of riparian disturbance used in the NRSA are themselves direct indicators of human disturbance, and were used to screen reference sites, we illustrate the relationship of W1_HALL to the human disturbance gradient in Figures 1 and 2 only to compare their relative magnitudes among disturbed and undisturbed streams in the various regions of the U.S.

7.6 Literature cited

Barbour, M. T., J. Gerritsen, B. D. Snyder, and J. B. Stribling. 1999. Rapid bioassessment protocols for use in streams and wadeable rivers: periphyton, benthic macroinvertebrates, and fish. EPA/841-B-99-002, U.S. Environmental Protection Agency, Washington, D.C.

Benson, B. J. and J. J. Magnuson. 1992. Spatial heterogeneity of littoral fish assemblages in lakes: relation to species diversity and habitat structure. Canadian Journal of Fisheries and Aquatic Sciences 49:1493-1500.

Bisson, P. A., J. L. Nielsen, R. A. Palmason, and L. E. Grove. 1982. A system of naming habitat types in small streams, with examples of habitat utilization by salmonids during low stream flow. Pages 62-73 In N. B. Armantrout, [editor]. Acquisition and utilization of aquatic habitat inventory information. Symposium Proceedings, October 28-30, 1981, Portland, Oregon. The Hague Publishing, Billings, Montana.

Bryce, S.A., G.A. Lomnicky, P.R. Kaufmann, L.S. McAllister, and T.L. Ernst. 2008. Development of Biologically-Based Sediment Criteria in Mountain Streams of the Western United States. N. Am. J. Fish. Manage. 28: 28:1714-1724.

Bryce, S.A., Lomnicky, G.A, and P.R. Kaufmann. 2010. Protecting Sediment-Sensitive Aquatic Species in Mountain Streams through the Application of Biologically-Based Criteria Streambed Sediment Criteria. J. North American Benthological Soc. 29(2):657-672.

Buffington, J. M. and D. R. Montgomery. 1999a. Effects of hydraulic roughness on surface textures of gravel-bed rivers. Water Resources Research 35:3507-3521.

Buffington, J. M. and D. R. Montgomery. 1999b. Effects of sediment supply on surface textures of gravel-bed rivers. Water Resources Research 35:3523-3530.

Cummins, K. W. 1974. Structure and function of stream ecosystems. BioScience 24:631-641.

Dietrich, W. E., J. W. Kirchner, H. Ikeda, and F. Iseya. 1989. Sediment supply and the development of the coarse surface layer in gravel bed rivers. Nature 340:215-217.

Dingman, S. L. 1984. Fluvial Hydrology. W.H. Freeman, New York.

Dunne, T. and L. B. Leopold. 1978. Water in environmental planning. W. H. Freeman and Co., New York

Faustini, J. M. and P. R. Kaufmann. 2007. Adequacy of Visually Classified Particle Count Statistics From Regional Stream Habitat Surveys1. Journal of the American Water Resources Association 43:1293-1315.

Fitzpatrick, F. A., I. R. Waite, P. J. D'Arconte, M. R. Meador, M. A. Maupin, and M. E. Gurtz. 1998. Revised methods for characterizing stream habitat in the National Water-Quality Assessment Program. Water-Resources Investigations Report 98-4052, U.S. Geological Survey Reston, Virginia.

Frissell, C. A., W. J. Liss, C. E. Warren, and M. D. Hurley. 1986. A hierarchical framework for stream habitat classification: viewing streams in a watershed context. Environmental Management 10:199-214.

Gordon, N. D., T. A. McMahon, and B. L. Finlayson. 1992. Stream hydrology, an introduction for ecologists. John Wiley & Sons, New York.

Gorman, O. T. and J. R. Karr. 1978. Habitat structure and stream fish communities. Ecology 59:507-515.

Gregory, S. V., F. J. Swanson, W. A. McKee, and K. W. Cummins. 1991. An Ecosystem Perspective of Riparian Zones. BioScience 41:540-551.

Hankin, D. G. and G. H. Reeves. 1988. Estimating total fish abundance and total habitat area in small streams based on visual estimation methods. Canadian Journal of Fisheries and Aquatic Sciences 45:834-844.

Harmon, M. E., J. F. Franklin, F. J. Swanson, P. Sollins, S. V. Gregory, J. D. Lattin, N. H. Anderson, S. P. Cline, N. G. Aumen, J. R. Sedell, G. W. Lienkamper, K. Cormack Jr., and K. W. Cummins. 1986. Ecology of Coarse Woody Debris in Temperate Ecosystems. Advances in Ecological Research 15:133-302.

Harrelson, C. C., C. L. Rawlins, and J. P. Potyondy. 1994. Stream channel reference sites: an illustrated guide to field technique. General Tech. Rep. RM-245, USDA Forest Service, Rocky Mountain Forest and Range Experiment Station, Fort Collins, Colorado.

Hawkins, C. P., J. L. Kershner, P. A. Bisson, M. D. Bryant, L. M. Decker, S. V. Gregory, D. A. McCullough, C. K. Overton, G. H. Reeves, R. J. Steedman, and M. K. Young. 1993. A Hierarchical Approach to Classifying Stream Habitat Features. Fisheries 18:3-12.

Hawkins, C. P., M. L. Murphy, and N. H. Anderson. 1983. Density of fish and salamanders in relation to riparian canopy and physical habitat in streams of the northwestern United States. Canadian Journal of Fisheries and Aquatic Sciences 40:1173-1186.

Herlihy, A.T., S.G. Paulsen., J. Van Sickle, J.L. Stoddard. C.P. Hawkins, L.L. Yuan. 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(4):860–877.

Hynes, H. B. N. 1970. The ecology of running waters. Unversity of Toronto Press, Toronto, Ontario, Canada.

Kappesser, G. B. 2002. A riffle stability index to evaluate sediment loading to streams. Journal of the American Water Resources Association 38:1069-1081.

Kaufmann, P. R. 1987a. Channel morphology and hydraulic characteristics of torrent-impacted forest streams in the Oregon Coast Range, U.S.A. Ph.D. Dissertation. Oregon State University, Corvallis, Oregon.

Kaufmann, P. R. 1987b. Slackwater habitat in torrent-impacted streams. Pages 407-408 In R. L. Beschta, T. Blinn, G. E. Grant, F. J. Swanson, and G. E. Ice, [editors]. Erosion and Sedimentation in the Pacific Rim. International Association of Hydrologic Science, Pub. No. 165, Proceedings of an International Symposium, August 3-7, 1986, Ore. State Univ., Corvallis, OR. International Association of Hydrologic Science.

Kaufmann, P. R. 1993. Physical habitat. Pages 59-69 In R. M. Hughes. Stream Indicator and Design Workshop. EPA 600/R-93/138, U.S. Environmental Protection Agency, Office of Research and Development, Corvallis, Oregon.

Kaufmann, P. R., J. M. Faustini, D. P. Larsen, and M. A. Shirazi. 2008. A roughness-corrected index of relative bed stability for regional stream surveys. Geomorphology 199:150-170.

Kaufmann, P.R., D.P. Larsen, and J.M. Faustini, 2009. Bed Stability and Sedimentation Associated With Human Disturbances in Pacific Northwest Streams. J. Am. Water Resources Assoc. 45(2):434-459.

Kaufmann, P. R., P. Levine, E. G. Robison, C. Seeliger, and D. V. Peck. 1999. Quantifying physical habitat in wadeable streams. EPA/620/R-99/003, U.S. Environmental Protection Agency, Washington, D.C.

Kaufmann, P. R. and T. R. Whittier. 1997. Habitat Assessment. Pages 5-1 to 5-26 In J. R. Baker, D. V. Peck, and D. W. Sutton. Environmental Monitoring and Assessment Program -Surface Waters: Field Operations Manual for Lakes. EPA/620/R-97/001, U.S. Environmental Protection Agency, Washington, D.C.

Kaufmann, P. R. & J. M. Faustini, 2011. Simple measures of channel habitat complexity predict transient hydraulic storage in streams. Hydrobiologia. doi:10.1007/s10750-011-0841-y.

Keim, R. F., A. E. Skaugset & D. S. Bateman, 2002. Physical aquatic habitat II. Pools and cover affected by large woody debris in three western Oregon streams. North American Journal of Fisheries Management 22: 151–164.

Kovalenko, K.E., S.M. Thomaz, and D.M. Warfe. 2012. Habitat complexity: approaches and future directions – editorial review. Hydrobiologia 685:1–17. DOI 10.1007/s10750-011-0974-z

Larsen, D. P., P. R. Kaufmann, T. M. Kincaid, and N. S. Urquhart. 2004. Detecting persistent change in the habitat of salmon-bearing streams in the Pacific Northwest. Canadian Journal of Fisheries and Aquatic Sciences 61:283-291.

Lazorchak, J. M., D. J. Klemm, and D. V. e. Peck. 1998. Environmental Monitoring and Assessment Program-Surface Waters: field operations and methods for measuring the ecological condition of wadeable streams. EPA/620/R-94/004F, U.S. Environmental Protection Agency, Washington, D.C.

Leopold, L. B. 1994. A View of the River. Harvard University Press, Cambridge, Massachusetts.

Leopold, L. B., M. G. Wolman, and J. P. Miller. 1964. Fluvial processes in geomorphology. W.H. Freeman, San Francisco.

Lisle, T. E. 1982. Effects of aggradation and degradation on riffle-pool morphology in natural gravel channels, northwestern California. Water Resources Research 18:643-1651.

Lisle, T. E. 1987. Using “residual depths” to monitor pool depths independently of discharge. Research Note PSW-394, USDA Forest Service, Pacific Southwest Forest and Range Experimental Station, Berkeley, California.

Lisle, T. E. and S. Hilton. 1992. The volume of fine sediment in pools: an index of sediment supply in gravel-bed streams. Water Resources Bulletin 28:371-383.

MacDonald, L. H., A. W. Smart, and R. C. Wismar. 1991. Monitoring Guidelines to Evaluate Effects of Forestry Activities on Streams in the Pacific Northwest and Alaska. EPA 910/9-91-001, U.S. Environmental Protection Agency, Region X, Seattle, Washington.

Madej, M. A. 2001. Development of channel organization and roughness following sediment pulses in single-thread, gravel bed rivers. Water Resources Research 37:2259-2272.

Montgomery, D. R. and J. M. Buffington. 1993. Channel classification, prediction of channel response, and assessment of channel condition. Washington State Timber/Fish/Wildlife Agreement, Report TFW-SH10-93-002, Department of Natural Resources, Olympia, Washington.

Montgomery, D. R. and J. M. Buffington. 1997. Channel-reach morphology in mountain drainage basins. Geological Society of America Bulletin 109.

Montgomery, D. R. and J. M. Buffington. 1998. Channel processes, classification, and response. Pages 13-42 In R. Naiman and R. Bilby, [editors]. River Ecology and Management. Springer-Verlag, New York.

Moore, K. M. S. and S. V. Gregory. 1988. Summer habitat utilization and ecology of cutthroat trout fry (*Salmo clarki*) in Cascade mountain streams. Canadian Journal of Fisheries and Aquatic Sciences 45:1921-1930.

Morisawa, M. 1968. Streams, their dynamics and morphology. McGraw-Hill Book Company, New York.

Mossop, B. & M. J. Bradford, 2006. Using thalweg profiling to assess and monitor juvenile salmon (*Oncorhynchus spp*.) habitat in small streams. Canadian Journal of Fisheries and Aquatic Sciences 63: 1515–1525.

Naiman, R. J., H. Decamps, J. Pastor, and C. A. Johnston. 1988. The potential importance of boundaries to fluvial ecosystems. Journal of the North American Benthological Society 7:289-306.

O'Neill, M. P. and A. D. Abrahams. 1984. Objective identification of pools and riffles. Water Resources Research 20:921-926.

Paulsen, S.G., A. Mayio, D.V. Peck, J.L. Stoddard, E.Tarquinio, S.M. Holdsworth, J. Van Sickle, L.L. Yuan, C.P. Hawkins, A.T. Herlihy, P.R. Kaufmann, M.T. Barbour, D.P. Larsen, and A.R. Olsen. 2008. Condition of stream ecosystems in the US: an overview of the first national assessment. J. N. Am. Benthological Soc. 27(4):812–821.

Pearsons, T. N. & G. M. Temple, 2007. Impacts of early stages of salmon supplementation and reintroduction programs on three trout species. North American Journal of Fisheries Management 27: 1–20.

Pearsons, T. N. & G. M. Temple, 2010. Changes to Rainbow Trout abundance and salmonid biomass in a Washington watershed as related to hatchery salmon supplementation. Transactions of the American Fisheries Society 139: 502–520.

Peck, D. V., D. K. Averill, A. T. Herlihy, B. H. Hill, R. M. Hughes, P. R. Kaufmann, D. J. Klemm, J. M. Lazorchak, F. H. McCormick, S. A. Peterson, P. L. Ringold, M. R. Cappaert, T. Magee, and P. A. Monaco. in press. Environmental Monitoring and Assessment Program: Surface Waters Western Pilot Study—field operations manual for nonwadeable streams. EPA 620/ R-xx/xxx, U.S. Environmental Protection Agency, Washington, D.C.

Peck, D. V., A. T. Herlihy, B. H. Hill, R. M. Hughes, P. R. Kaufmann, D. J. Klemm, J. M. Lazorchak, F. H. McCormick, S. A. Peterson, P. L. Ringold, T. Magee, and M. R. Cappaert. 2006. Environmental Monitoring and Assessment Program: Surface Waters Western Pilot Study—field operations manual for wadeable streams. EPA/620/R-06/003, U.S. Environmental Protection Agency, Washington, D.C.

Platts, W. S., W. F. Megahan, and G. W. Minshall. 1983. Methods for evaluating stream, riparian, and biotic conditions. General Technical Report INT-138, U.S. Department of Agriculture, Forest Service, Intermountain Forest and Range Experiment Station, Ogden, Utah.

Rinne, J. 1988. Effects of livestock grazing exclosure on aquatic macroinvertebrates in a montane stream, New Mexico. Great Basin Naturalist 48:146-153.

Robison, E. G. and R. L. Beschta. 1989. Estimating stream cross sectional area from wetted width and thalweg depth. Physical Geography 10:190-198.

Robison, E. G. and P. R. Kaufmann. 1994. Evaluating two objective techniques to define pools in small streams. Pages 659-668 In R. A. Marston and V. A. Hasfurther, [editors]. Effects of Human Induced Changes on Hydrologic Systems. Summer Symposium proceedings, American Water Resources Association,. June 26-29, 1994, Jackson Hole, Wyo.

Stack, W. R. and R. L. Beschta. 1989. Factors influencing pool morphology in Oregon coastal streams. Pages 401-411 In W. W. Woessner and D. F. Potts, [editors]. Headwaters Hydrology Symposium. American Water Resources Association.

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

Stoddard, J. L., D. V. Peck, S. G. Paulsen, J. Van Sickle, C. P. Hawkins, A. T. Herlihy, R. M. Hughes, P. R. Kaufmann, D. P. Larsen, G. Lomnicky, A. R. Olsen, S. A. Peterson, P. L. Ringold, and T. R. Whittier. 2005b. An ecological assessment of western streams and rivers. EPA 620/R-05/005, U.S. Environmental Protection Agency, Washington, D.C.

Suttle, K. B., M. E. Power, J. M. Levine, and C. McNeely. 2004. How fine sediment in riverbeds impairs growth and survival of juvenile salmonids. Ecological Applications 14:969–974.

U.S. EPA. 2004. Wadeable Streams Assessment: field operations manual. EPA/841/B-04/004, U.S. Environmental Protection Agency, Washington, D.C.

U.S. EPA. 2006. Wadeable Streams Assessment: a collaborative survey of the Nation's streams. EPA/641/B-06/002, U.S. Environmental Protection Agency, Washington, D.C.

U.S. EPA 2007. National Rivers and Streams Assessment: Field Operations Manual EPA-841-B-07-009. U.S. Environmental Protection Agency, Washington, D.C.

Wilcock, P. R. 1997. The components of fractional transport rate. Water Resources Research 33:247-258.

Wilcock, P. R. 1998. Two-fraction model of initial sediment motion in gravel-bed rivers. Science 280:410-412.

Table G-1. Metrics used to characterize the general attributes of stream and river physical habitat.

Habitat Volume:
- *LRP100 = log(RP100)* = Log of Mean Residual Depth (cm)

Scaled Habitat Volume:
- *LDVRP100 = log(RP100) – log(Predicted RP100)* = Deviation in Mean Residual Depth from expected value

Habitat Complexity:
- *CVDPTH = SDDEPTH / XDEPTH* = Coefficient of Thalweg Depth Variation
- *C1WM100* = Number of Large Woody Debris pieces/100m of channel.
- *LV1W_MSQ* = log[Volume of Large Woody Debris per m^2 of bankfull channel area (m^3/m^2)].
- *XFC_NAT* = Areal Cover of Woody Debris, Brush, Undercut Banks, Overhanging Vegetation, plus Boulders and Rock Ledges.
- *XFC_NORK* = Areal Cover of Woody Debris, Brush, Undercut Banks, Overhanging Veg.
- *XFC_AQM* = Areal Cover of Aquatic Macrophytes
- *XFC_ALG* = Areal Cover of Filamentous Algae detectable by the unaided eye.

Streambed Particle Size:
- *LSUB_dmm* = log[Streambed surface particle D_g – mm] = log of geometric mean diameter of bed surface sediments in millimeters.
- *PCT_FN* = % Streambed Silt & Finer
- *PCT_SAFN* = % Streambed Sand & Finer
- *XEMBED* = % Substrate Embedded by Sand and Fines

Scaled Streambed Particle Size:
- *DPCT_FN* = Deviation of *PCT_FN* from expected value ("excess Fines")
- *DPCT_SF* = Deviation of *PCT_SAFN* from expected value ("excess Sand+Fines")
- *DEVLSUB* = Deviation of *LSUB_DMM* from expected value (Streambed Fining Index)

Relative Bed Stability:
- *LRBS*= $\log_{10}$ of diameter ratio: Geometric mean bed particle diameter / Critical (mobile) diameter at bankfull flow stage. (LRBS_bw5: see Kaufmann et al. 1999; LRBS_g08: see Kaufmann et al. 2008, 2009).

Table G-1. (Continued). Metrics used to characterize the general attributes of stream and river physical habitat.

Floodplain Interaction:

- *LSINU = Log(SINU)* = Log(Channel Sinuosity).
- *LINCIS_H = log(XINC_H - XBKF_H + 0.1)* = Log of Incision from terrace to bankfull ht (m).
- LBFWDRAT = log{BKF_W / BKF_H+(XDEPTH/100)}= log (Bankfull Width/Depth Ratio)
- *LBFXWRAT = log(BKF_W / XWIDTH)*= log (Bankfull Width / Wetted Width) (an index of streamside flood inundation potential)

Hydrologic Regime:

- LQSLTR_RAT = log{(Qsp+0.0000001)/LTROFF_M}=log{low flow /annual mean runoff} (~ an inverse index of "droughtiness", where: Qsp = Flow_mps/WSAREAKM= (flow_cfs/35.315)/WSAREAKM
- *LBFXDRAT =log{(XBKF_H+(XDEPTH/100) / (XDEPTH/100)}*= log(ratio of bankfull depth / wetted depth), a morphometric index of "flashiness".

Riparian Vegetation:

- *XCDENMID*: % Canopy Density measured midstream.
- *XCMG* = Riparian Canopy+Mid-+Ground Layer Vegetation (areal cover proportion)
- *XCMGW* = Riparian Canopy+Mid+Ground Layer <u>Woody</u> Veg.(areal cover proportion)

Riparian Habitat Alteration:

- $QR1=(QRVEG1*QRVEG2*QRDIST1)^{0.3333}$; where:
  if *XCMGW* <=2.00 then *QRVeg1*=.1+(0.9(*XCMGW* / 2.00));
  if XCMGW >2.00 then QRVeg1=1;
- QRVeg2=.1+(0.9(XCDENBK / 100)); and QRDIST1=1/(1+W1_HALL)

Riparian Human Disturbances:

- *W1_HAG* = Riparian & near-Stream Agriculture – all types (proximity-weighted tally)
- *W1H_ROAD* = Riparian & near-Stream Roads (proximity-weighted tally)
- *W1H_CROP* = Riparian & near-Stream Row Crop Agriculture (proximity-weighted tally)
- *W1H_WALL* = Riparian & near-Stream Walls, Dikes, Revetment (proximity-weighted tally)
- *W1_HALL* = Proximity-weighted Index of Human Disturbances of All Types
- *QRDIST1* = 1/(1+*W1_HALL*) = Proximity-weighted Inverse Index of Human Disturbances of All Types

Table G-2. Sampling revisit precision (repeatability) of the four physical habitat condition indicators selected for the National Rivers and Streams Assessment. Repeat visits within the summer sampling season were used to calculate RMS_{rep}, which is essentially the standard deviation of repeat sampling pairs to the same stream or river reach. Dividing the square of the RMS_{rep} into the variance among sites gives the S/N variance ratio. (See Kaufmann et al. 1999 for ANOVA methods to calculate RMS_{rep} and S/N, where RMS_{rep} is equal to their RMSE.)

Metric	Group	Sites (n)	mean	Repeat pairs (n)	RMS_{rep}	S/N
LRBS_g08	**All Sites**	**1945**	**-0.776**	**177**	**0.482**	**4.97**
	Boatable	711	-0.636	89	0.450	7.36
	Wadeable	1234	-0.860	88	0.512	3.31
	EHIGH	534	-0.397	70	0.500	4.19
	PLNLOW	1002	-1.014	74	0.494	4.67
	WMTNS	409	-0.712	33	0.411	6.07
L_xfc_nat	**All Sites**	**2113**	**-0.590**	**197**	**0.240**	**1.87**
	Boatable	782	-0.575	93	0.242	1.22
	Wadeable	1331	-0.599	104	0.238	2.29
	EHIGH	555	-0.460	73	0.209	0.92
	PLNLOW	1125	-0.675	86	0.263	1.78
	WMTNS	433	-0.545	38	0.241	1.77
L_xcmgw	**All Sites**	**2113**	**-0.286**	**197**	**0.146**	**9.38**
	Boatable	782	-0.175	93	0.155	4.28
	Wadeable	1331	-0.353	104	0.137	13.20
	EHIGH	555	-0.062	73	0.092	6.01
	PLNLOW	1125	-0.381	86	0.174	8.53
	WMTNS	433	-0.340	38	0.162	5.74
L_W1_Hall	All Sites	2113	-0.152	197	0.186	5.18
	Boatable	782	-0.123	93	0.145	7.99
	Wadeable	1331	-0.170	104	0.216	3.89
	EHIGH	555	-0.108	73	0.184	5.08
	PLNLOW	1125	-0.189	86	0.171	6.02
	WMTNS	433	-0.116	38	0.220	4.00

Table G-3. Number of years required for a 50-site monitoring network to detect 1% and 2% per year trends in habitat attributes with 80% likelihood (*beta,* or power) and *alpha* = 0.05, if specified trends occur, and sites are visited each year taken from Larsen et al. (2004),[a] or calculated using the same data and analytical procedures used in that publication.[b]

Variable		1% trend	2% trend
SDDEPTH[b]	(Std. Deviation of Thalweg Depth)	13 years	8 years
LRP100[a]	(log[Mean Residual Depth])	20	12
PCT_SAFN[a]	(% Sand + Silt)	21	13
XEMBED[b]	(% Embeddedness)	20	12
LRBS_BW5[b]	(log[Rel. Bed Stability])	12	8
LV1W_MSQ[a]	(log[Large Wood Volume/m^2])	27	17
XCMGW[b]	(3-Layer Riparian Woody Veg Cover)	12	8
XCDENMID[a]	(Canopy Density)	13	8

Table G-4. Anthropogenic disturbance screening criteria used in the National Rivers and Streams Assessment to characterize least-disturbed reference (R), moderately disturbed (M), and most-disturbed (D) sample reaches for developing physical habitat condition criteria. Values > than those before the slash (/) are EXCLUSION criteria for reference sites. Values ≥ those after slash are INCLUSION criteria for most-disturbed sites.

Region	PTL	NTL	Cl	SO4	Turb	W1_HALL	W1_HAG Wadeable	W1H_CROP Wadeable	W1H_WALL Wadeable	PCTCROP	PCTURB	DamScreen
NAP	20/100	750/3500	250/10000	250/1000	5/10	2.0/4.0	0.1/0.4	0.05/0.10	0.2/0.4	15/67	5/25	1/1
SAP	20/100	750/3500	200/1000	400/1000	5/20	2.0/4.0	0.1/0.4	0.05/0.10	0.2/0.4	15/67	5/25	1/1
UMW	50/150	1000/5000	300/2000	400/2000	5/30	2.0/4.0	0.15/1.4	0.1/0.4	0.2/0.4	15/67	5/25	1/1
CPL	75/250	2500/8000	999999/ 999999	600/4000	10/50	2.0/4.0	0.15/1.4	0.05/ 0.4	0.2/0.4	15/67	5/25	1/1
TPL	100/500	3000/15000	2000/5000	999999/ 999999	50/100	2.0/4.0	0.67/1.4	0.25/0.48	0.4/0.6	15/67	5/25	1/1
SPL	150/500	4500/10000	1000/5000	999999/ 999999	50/100	2.0/3.0	1.0/1.4	0.15/ 0.25	0.2/0.4	15/67	5/25	1/1
WMT:												
Southwest	50/100	750/1500	300/1000	99999/ 99999	5/10	W:0.5/3.0 B,G:1.5/3.0	0.25/1.4	0.10/0.25	0.2/0.4	15/67	5/25	1/1
S.Rockies	25/100	750/1500	200/1000	200/1000	5/10	W:1.0/3.0 B,G:1.5/3.0	0.3/1.4	0.1/0.25	0.2/0.4	15/67	5/25	1/1
N.Rockies Pacific	25/100	750/1500	200/1000	200/1000	5/10	W:0.5/3.0 B,G:1.5/3.0	0.3/1.4	0.10/0.25	0.2/0.4	15/67	5/25	1/1
XER	50/150	1500/5000	1000/5000	999999/ 999999	25/75	1.5/3.0	0.6/1.4	0.15/0.25	0.2/0.4	15/67	5/25	1/1

Table G-5. Responsiveness of NRSA physical habitat condition metrics to levels of human disturbance, as quantified by *t*-values of the difference between means of least-disturbed reference sites (RMD_PHab=R) minus more-disturbed sites (those screened as RMD_PHab=D). Values shown in red have a sign contrary to expectations.

Metric	**Region**	**t-value R-D (Boatable)**	**t-value R-D (Wadeable)**	**t-value R-D (All sites)**
LRBS_g08	USA-48	+5.12	+5.82	+7.85
	CPL	+2.36	+0.37	+2.11
	EHIGH (NAP+SAP)	+3.03	+3.01	+4.29
	INTPLNUMW (NPL,SPL,TPL,UMW)	+3.09	+3.68	+4.65
	West (WMT+XER)	+3.19	+5.07	+5.21
LXFC_Nat_OE	USA-48	-3.02	+6.66	+4.25
	CPL	+0.18 (ns)	+3.78	+3.09
	EHIGH (NAP+SAP)	-1.79	-0.12 (ns)	-1.26
	INTPLNUMW (NPL,SPL,TPL,UMW)	-2.30	+3.19	+2.13
	West (WMT+XER)	-1.87	+6.94	+4.17
L_xcmgw	USA-48	+2.38	+12.61	+12.24
	CPL	+4.44	+4.71	+6.10
	EHIGH (NAP+SAP)	+0.08 (ns)	+5.19	+4.11
	INTPLNUMW (NPL,SPL,TPL,UMW)	+0.56 (ns)	+8.66	+8.24
	West (WMT+XER)	+2.16	+6.46	+6.85

Appendix G-1. Reference condition models

A) **Channel Bed Sedimentation based on Relative Bed Stability** (RBS) (LRBS_use=LRBS_g08 = Log10(RBS_g08) =calculated according to Kaufmann et al. (2008)

Following are Simple LRBS models (reference distributions) or regression models in which W1_Hall is set to zero in regressions on W1_Hall within RMD_PHAB=R --- Then mean and SD of adjusted ref mean LRBS becomes y-intercept and the SD about the reference mean is the RMSE of those regressions.

Coastal Plain (CPL) combined Boatable & Wadeable reference sites model:

LRBS= -0.9855 – 1.0320(W1_Hall)
R^2=0.2585, RMSE=0.7123, p=0.0003, df=45 (minus 3 hardpan Boatable sites with LRBS>2)

Condition classes (use y-int -0.67 x & 1.65x RMSE for 25th and 5th %-tiles of the reference distribution)
5th %-tile = - 2.161
25th %-tile= -1.463

Eastern Highlands (NAP & SAP) combined Boatable & Wadeable reference site model:

LRBS= -0.1977 + 0.3311(W1_Hall) --------- (note positive slope)
R^2=0.0357, RMSE=0.8445, p=0.0624, df= 97 (NO outliers removed; appears to be subclass of low RBS)

Condition classes (use y-int-0.67x & 1.65x RMSE for 25th and 5th %-tiles of the reference distribution)
5th %-tile = -1.591
25th %-tile= -0.764

Combined Upper Midwest plus Temperate and Southern Plains (UMW, TPL, SPL) combined Boatable & Wadeable reference site model:

LRBS= -0.3126 – 0.8593(W1_Hall)
R^2=0.0741, RMSE=1.2239, p=0.0052, df=103

Condition classes (use y-int-0.67x & 1.65x RMSE for 25th and 5th %-tiles of the reference distribution
5th %-tile = -2.332
25th %-tile= -1.133

Northern Plains (NPL) combined Boatable & Wadeable reference site model:

Simple Distrib: ref mean LRBS= -0.333 with SD=0.824, df=23 (1 low outlier removed)

Condition classes (use ref mean -0.67x & 1.65x SD for 25th and 5th %-tiles of the reference distribution)

5th %-tile = -1.692
25th %-tile= -0.885

<u>The West (WMT & XER) Separate Boatable and Wadeable reference site models</u>:

<u>Boatable Reference site model</u>:
LRBS= +0.5727 – 0.4064(W1_Hall)
R^2=0.0555, RMSE=0.6818, p=0.2677, df=23
(retained 2 low outliers given the small sample size)
(Weak model but scope of W1_HALL is small and the same, but stronger, relationship is observed across all sites.)

Condition classes (use y-int-0.67x & 1.65x RMSE for 25th and 5th %-tiles of the reference distribution)
5th %-tile = - 0.5523
25th %-tile= +0.1159

<u>Wadeable Reference site model</u>:
LRBS= -0.5207 -0.4818(W1_Hall)
R^2=0.0225, RMSE=0.7006, p=2486, df=60
(Weak model but scope of w1hall small and same, but stronger relationship across all sites.)

Condition classes (use y-int-0.67x & 1.65x RMSE for 25th and 5th %-iles of the reference distribution)
5th %-tile = -1.677
25th %-tile= -0.990

B) Instream Fish Cover XFC_NAT (Transformed as Log_{10}(0.01+XFC_NAT)

Following are aggregated region Observed/Expected (O/E) models with disturbance modeled out where possible 6/13/12;

<u>Coastal Plain (CPL) combined Boatable & Wadeable reference site model</u>:
(In Reference Sites, XFC_NAT increases slightly in with QR1 or RDIST1 -- over all sites the trend is strong but the ref sites are generally high in fish cover (no need to separate Wadeable and Boatable sites).

LXFC_NAT $_{ref}$ =-0.6471+0.3820(QR1)
R^2=0.0333 RMSE=0.2656 p=0.20, df=49

Expected reference values calculated by setting QR1 to 0.9 (90th percentile = 0.89)
LXFC_NAT_E = -0.6471 +(0.3820*0.9)
LXFC_NAT_OE= LXFC_NAT – LXFC_NAT_E

use mean zero and Emodel RMSE as above

LXFC_NAT_OE 25th % tile = -0.67 x 0.2656
LXFC_NAT_OE 5th % tile = -1.65 x 0.2656

Eastern Highlands (NAP & SAP) combined Boatable + Wadeable reference site model:

$LXFC\_NAT_{ref}$ = -0.3211 - 0.0348(Tmax_PT) +0.1873(L_ELEV_PTx) + 0.2545(RDIST1) + 0.5455(QR1)

R^2=0.2488, RMSE=0.2796, $p<0.0001$, df=99

significance values for model predictors:

Tmax_PT:	p=0.0006
L_ELEV_PTx	p=0.0223
RDIST1	p=0.1609
QR1	p=0.0325

Expected reference values calculated by setting QR1=0.9:

LXFC_NAT_E= -0.3211 -0.0348(Tmax_PT) +0.1873(L_ELEV_PTx) +0.2545(0) +0.5455(0.90)

mean zero and Emodel RMSE;

LXFC_NAT_OE 5th % tile = -1.65 x 0.2796
LXFC_NAT_OE 25th % tile = -0.67 x 0.2796

Interior Plains plus Upper Midwest (TPL, NPL, SPL, UMW) combined Boatable & Wadeable reference site model:

$LXFC\_NAT_{ref}$ = 0.7371 -0.0399(Tmax_PT) +0.0251(Tmin_PT) -0.0005(Psum_PTx) +0.7052(QR1)

R^2=0.2369, RMSE=0.3168, $p<0.0001$, df=128

significance values for model predictors:

Tmax_PT:	p=0.0153
Tmin_PT:	p=0.0051
PSum_PTx	p<0.0001
QR1	p=0.0018

Expected reference Values calculated by setting QR1 to 0.80 (Note that 0.80 is the 95th percentile of QR1 among Ref sites in 'INTPLNUMW')

LXFC_NAT_E= 0.7371 -0.0399(Tmax_PT) +0.0251(Tmin_PT) -0.0005(Psum_PTx) +0.7052(0.80)

Mean reference site LXFC_NAT_OE= -0.0923 with SD= 0.3701
but that SD has disturbance variance in it -- use mean 0.0 and E-model RMSE

LXFC_NAT_OE 5th % tile = -1.65 x 0.3168
LXFC_NAT_OE 25th % tile = -0.67 x 0.3168

The West (WMT+XER) combined Boatable & Wadeable reference site model):

$LXFC\_NAT_{ref}$ = -1.10469 +0.0153(LAT_DD83) -0.0166(Tmin_PT) -0.5236(P_REALM) - 0.5817(W1_Hall) +1.1400(RDIST1)

R^2=0.5699, RMSE=0.2387, p<0.0001, df=81

significance values for model predictors:

LAT_DD83	p=0.0291
Tmin_PT:	p=0.0015
P_REALM	p<0.0001
W1_Hall)	p=0.0360
RDIST1	p=0.0421
QR1	p=0.0325

Expected reference values calculated by setting W1_Hall=0 and RDIST1=0;

LXFC_NAT_E= -1.10469 +0.0153(LAT_DD83) -0.0166(Tmin_PT) -0.5236(P_REALM) - 0.5817(0) +1.1400(0)

Mean reference site LXFC_NAT_OE = 0.0353 with SD= 0.2683

but that SD has disturbance variance in it -- use E-model RMSE and use mean 0.0.

LXFC_NAT_OE 5^{th} % tile = -1.65 x 0.2387

LXFC_NAT_OE 25^{th} % tile= -0.67 x 0.2387

Condition Class Thresholds for XFC_NAT:

Modeled 25^{th} and 5^{th} percentiles of the reference distribution:

$LXFC\_NAT\_OE_{ref}$ 5^{th} % tile = Mean $LXFC\_NAT\_OE_{ref}$ -1.65(SD_{ref})

$LXFC\_NAT\_OE_{ref}$ 25^{th} % tile = Mean $LXFC\_NAT\_OE_{ref}$ -0.67(SD_{ref});

Condition Criteria for LXFC_NAT:

If LXFC_NAT_OE < $LXFC\_NAT\_OE_{ref}$ 5^{th} % tile then FCvrCOND = 'P'

If LXFC_NAT_OE >= $LXFC\_NAT\_OE_{ref}$ 5^{th} % tile and LXFC_NAT_OE < $LXFC\_NAT\_OE_{ref}$ 25^{th} % tile then FCvrCOND='M'

If LXFC_NAT_OE >= $LXFC\_NAT\_OE_{ref}$ 25^{th} % tile then FCvrCOND='G';

If LXFC_NAT_OE =. then FCvrCOND='Z';

Where:

PSUM_PTx=PsumPY_PT;

IF PsumPY_PT=. then Psum_PTx=PSUM_WSx;

Define RDIST1 and QR1:

We calculated a composite riparian condition index (QR1) from the reach summary data describing the cover and structure of riparian vegetation and a proximity-weighted tally of streamside human activities. QR1 has a theoretical minimum approaching zero where there is no riparian vegetation and very high values of *W1_HALL,* the proximity weighted tally of streamside human land use activities. It approaches 1.0 where there is abundant, complex riparian woody vegetation, high bankside canopy density (measured with densiometer), and no visible human land use activities or channel alterations. It is intended for use in those riparian settings in regions where reference condition is a multi-storied woody vegetation corridor (XCMGW approaching 2.0), with bankside canopy density (XCDENBK) generally complete (85%-100%) along stream banks, and along rivers above bankfull height. Reference condition is set near zero for the types of riparian human activities identified by the EMAP Physical Habitat field methods (Peck et al. In Press-a; Peck et al., In Press-b). QR1 is then defined as the geometric mean of three scaled variables as follows (the cube-root is taken to reduce extreme skewness in the product of the three component variables:

QR1= $\{(QRVEG1)\ (QRVEG2)\ (QRDIST1)\}^{0.333}$; where:

if XCMGW <=2.00, then QRVeg1=.1+(.9 (XCMGW/2.00)), and
if XCMGW >2.00 then QRVeg1=1; and
QRVeg2=0.1 + [0.9(XCDENBK/100)];
QRDIST1=1/ (1+W1_HALL);
and W1_HALL= distance weighted tally of in-channel, riparian, and near stream human activities.

QR1 decreases with increases in streamside human activities (W1_HALL), and increases with increasing riparian woody vegetation complexity (XCMGW) and riparian cover density measured at the streambank with a canopy densiometer (XCDENBK).

We transformed the variable *W1_HALL,* a proximity-weighted tally of all the targeted types of human activities into an index that is more sensitive at the low end of disturbance and has a range constrained from 0 to 1. The new variable, QRDIST1=1/ (1+W1_HALL), is an *inverse* measure of riparian disturbance, with value of 1 when there are no observable human disturbances, and approaches 0 as the number and extent of human disturbances increases. Note that QRDIST1 was one of the component metrics used to define the riparian vegetation alteration variable QR1 in the previous section.

C) Riparian condition XCMGW (Transformed as L_xcmgw= Log_{10}(0.01+XCMGW):

Following are Simple L_xcmgw models (reference distributions) or regression models in which W1_Hall is set to zero in regressions on W1_Hall within RMD_PHAB=R --- The adjusted mean L_xcmgw for reference sites is defined as the y-intercept and the SD about the reference mean is defined as the RMSE of those regressions. Note that the primary reason for excluding outliers is to avoid gross overestimations of the reference SD.

<u>Coastal Plain (CPL) combined Boatable & Wadeable Reference Site model</u>:

L_xcmgw_{ref}= 0.0173 + 0.0846(W1_Hall)
(note here the W1_Hall slope is POSITIVE; its effect is to prevent overestimating the reference mean and it's RMSE (or SD)

R^2=0.0599, RMSE=0.1342, p=0.0835, df= 50

adjusted ref mean = 0.0173 (xcmgw=1.04)
SD=RMSE= 0.1342
Est 5th %-tile = ref mean x 1.65(SD) = -0.204 (xcmgw=0.625)
Est 25th %-tile = ref mean x 0.67(SD) = -0.0726 (xcmgw=0.846)

Combined Northern and Southern Appalachians (NAP&SAP):

NAP&SAP Boatable Reference Site Model:
L_xcmgw_{ref}= 0.0456 -0.0138(W1Hall) --- virtually the same as the null (simple) model.
R^2=0.0036, RMSE=0.1224, p=0.7347, df = 33 (excludes 3 NAP and 1 NAP low outliers)
null model: Mean L_xcmgwref=0.0334 and SD=0.1207, n=34 (excludes same 4 outliers)

adjusted ref mean =0.0456 (xcmgw=1.11)
SD=RMSE=0.1224
Est 5th %-tile = ref mean x 1.65(SD) = -0.156 (xcmgw=0.698)
Est 25th %-tile = ref mean x 0.67(SD) = -0.0364 (xcmgw=0.920)

NAP&SAP Wadeable Reference Site Model:
L_xcmgw_{ref}= 0.0823 -0.2064(W1Hall)
R^2 = 0.2490, RMSE=0.1284, p<0.0001, df = 61

adjusted ref mean =0.0823 (xcmgw=1.21)
SD=RMSE=0.1284
Est 5th %-tile = ref mean x 1.65(SD) = -0.130 (xcmgw=0.742)
Est 25th %-tile = ref mean x 0.67(SD) = -0.00373 (xcmgw=0.991)

Southern Plains (SPL) combined Boatable & Wadeable Reference Site Model:
L_xcmgw_{ref}= -0.3475 +0.2271(W1Hall)
(note here the W1_Hall slope is POSITIVE; its effect is to prevent overestimating the ref mean and it's RMSE (or SD)

R^2=0.1842, RMSE=0.2565, p=0.018, df = 29 (2 very low outliers removed)

adjusted ref mean = -0.3475 (xcmgw= 0.449)
SD=RMSE= 0.2565
Est 5th %-tile = ref mean x 1.65(SD) = -0.771 (xcmgw=0.170)
Est 25th %-tile = ref mean x 0.67(SD) = -0.519 (xcmgw=0.302)

Combined Upper Midwest, Northern Plains and Temperate Plains (UMW, NPL TPL):

UMW,NPL, &TPL Boatable reference site model:
L_xcmgw_{ref}= -0.0526 – 0.2840(W1_Hall)

R^2= 0.2098, RMSE=0.2664, p=0.0125, df = 28 (excludes 2 low outliers)
adjusted ref mean= -0.0526 (xcmgw= 0.886)
SD=RMSE=0.2664
Est 5th %-tile = ref mean x 1.65(SD) = -0.492 (xcmgw=0.322)
Est 25^{th} %-tile = ref mean x 0.67(SD) = -0.231 (xcmgw=0.587)

<u>UMW,NPL, &TPL Wadeable reference site model</u>:
L_xcmgw_{ref}= -0.1210 - 0.3276(W1_Hall)
R^2=0.2016, RMSE=0.2607, p<0.0001, df= 67 (excludes 2 low outliers)
adjusted ref mean= -0.1210 (xcmgw=0.757)
SD=RMSE=0.2607
Est 5th %-tile = ref mean x 1.65(SD) = -0.551 (xcmgw=0.281)
Est 25^{th} %-tile = ref mean x 0.67(SD) = -0.296 (xcmgw=0.506)

<u>Western Mountain (WMT) combined Boatable & Wadeable reference site model</u>:
L_xcmgw_{ref}= -0.1033 - 0.1586(W1_Hall)
R^2= 0.0716, RMSE=0.2356, p=0.0461, df= 55

adjusted ref mean = -0.1033 (xcmgw=0.788)
SD=RMSE=0.2356
Est 5th %-tile = ref mean x 1.65(SD) = -0.492 (xcmgw=0.322)
Est 25^{th} %-tile = ref mean x 0.67(SD) = -0.261 (xcmgw= 0.548)

<u>Xeric region (XER) combined Boatable & Wadeable reference site model</u>
L_xcmgw_{ref}= -0.1006 – 0.2467(W1_Hall)
R^2=0.1528, RMSE=0.2083, p=0.0397, df= 27 (1 low outlier removed)

adjusted ref mean = -0.1006 (xcmgw=0.793)
SD=RMSE=0.2083
Est 5th %-tile = ref mean x 1.65(SD) = -0.444 (xcmgw=0.360)
Est 25^{th} %-tile = ref mean x 0.67(SD) = -0.240 (xcmgw=0.575

Condition Class Thresholds for XCMGW:
Modeled 25^{th} and 5^{th} percentiles of the reference distribution:
L_xcmgw_{ref} 5^{th} % tile = Mean L_xcmgw_{ref} -1.65(SD_{ref})
L_xcmgw_{ref} 25^{th} % tile = Mean L_xcmgw_{ref} -0.67(SD_{ref});

Condition Criteria for XCMGW:
If L_xcmgw < $Lxcmgw_{ref}$ 5^{th} % tile then RIPCOND = 'P'
If L_xcmgw >= $Lxcmgw_{ref}$ 5^{th} % tile and L_xcmgw < $Lxcmgw_{ref}$ 25^{th} % tile then RIPCOND='M'
If L_xcmgw >= $Lxcmgw_{ref}$ 25^{th} % tile then RIPCOND='G';
If L_xcmgw =. then RIPCOND_rgrfw2='Z'

D) Riparian Human Disturbances (W1_Hall):

Uniform condition thresholds used nationwide:

L M X Means Low Medium and High DISTURBANCE --- this is analogous to the Good, Fair, Poor condition classification used for other metrics.

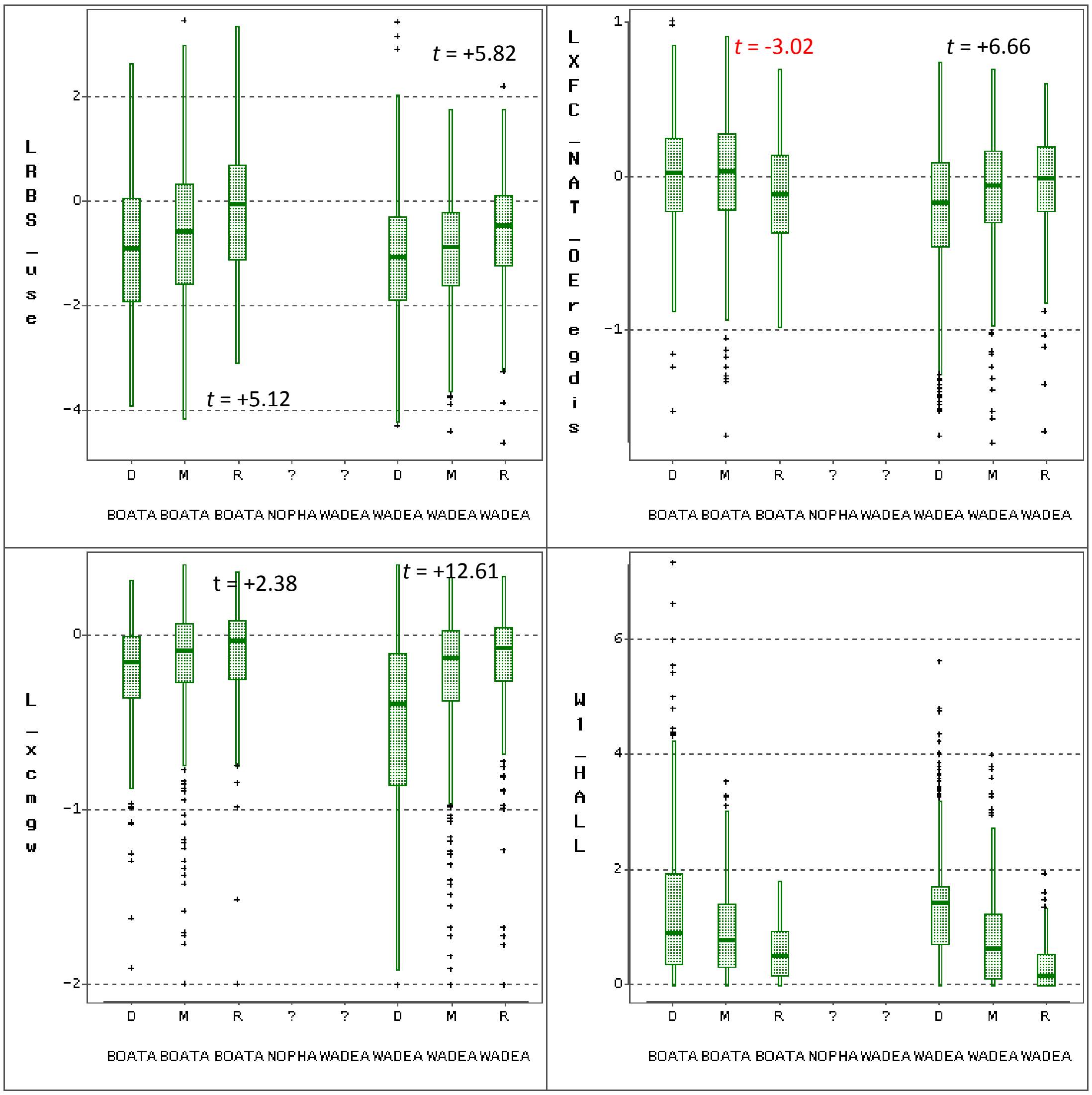
If W1_Hall < 0.33 then RDIST= 'L';

If W1_Hall >= 0.33 and W1_Hall < 1.5 then RDIST= 'M';

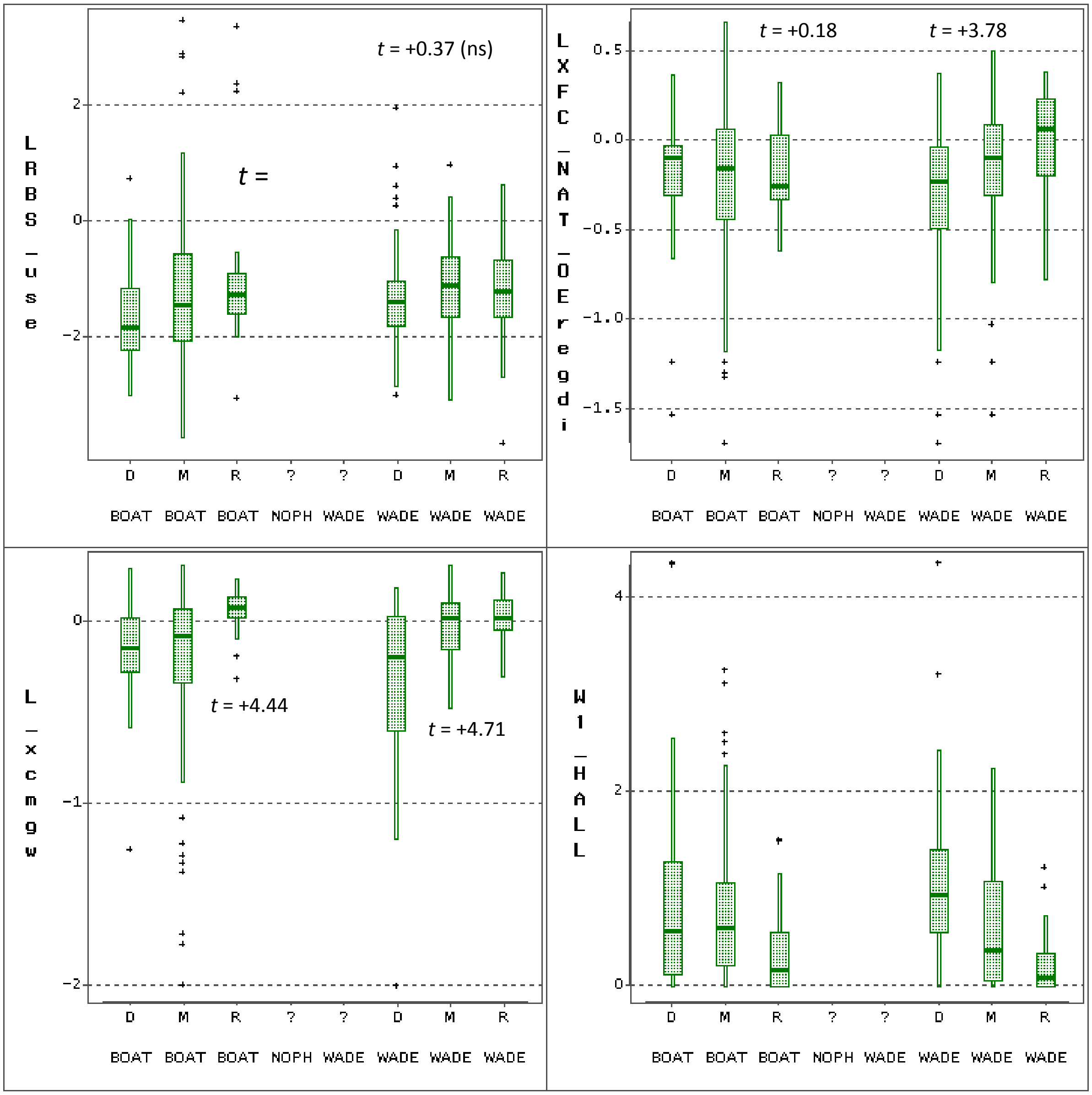
If W1_Hall >= 1.5 then RDIST = 'X';

If W1_Hall=. then RDIST = 'Z';

Figures G-1 A through J. Indicator response figures nationally and by the four combined regions used in Table G-5. LRBS_use = LRBS_g08, the indicator of relative bed stability and excess streambed fine sediments. LXCF_Nat_OEregdis = Log10 of observed/expected XFC_NAT, the indicator of instream habitat cover complexity. L_xcmgw = Log10(XCMGW), the indicator of riparian vegetation cover and structure. W1_HALL is the proximity weighted indicator of riparian and near-shore human disturbance intensity. Plots show *t* values for the differences between means for reference (R) and disturbed (D) sites. Values shown in red have a sign contrary to expectations.

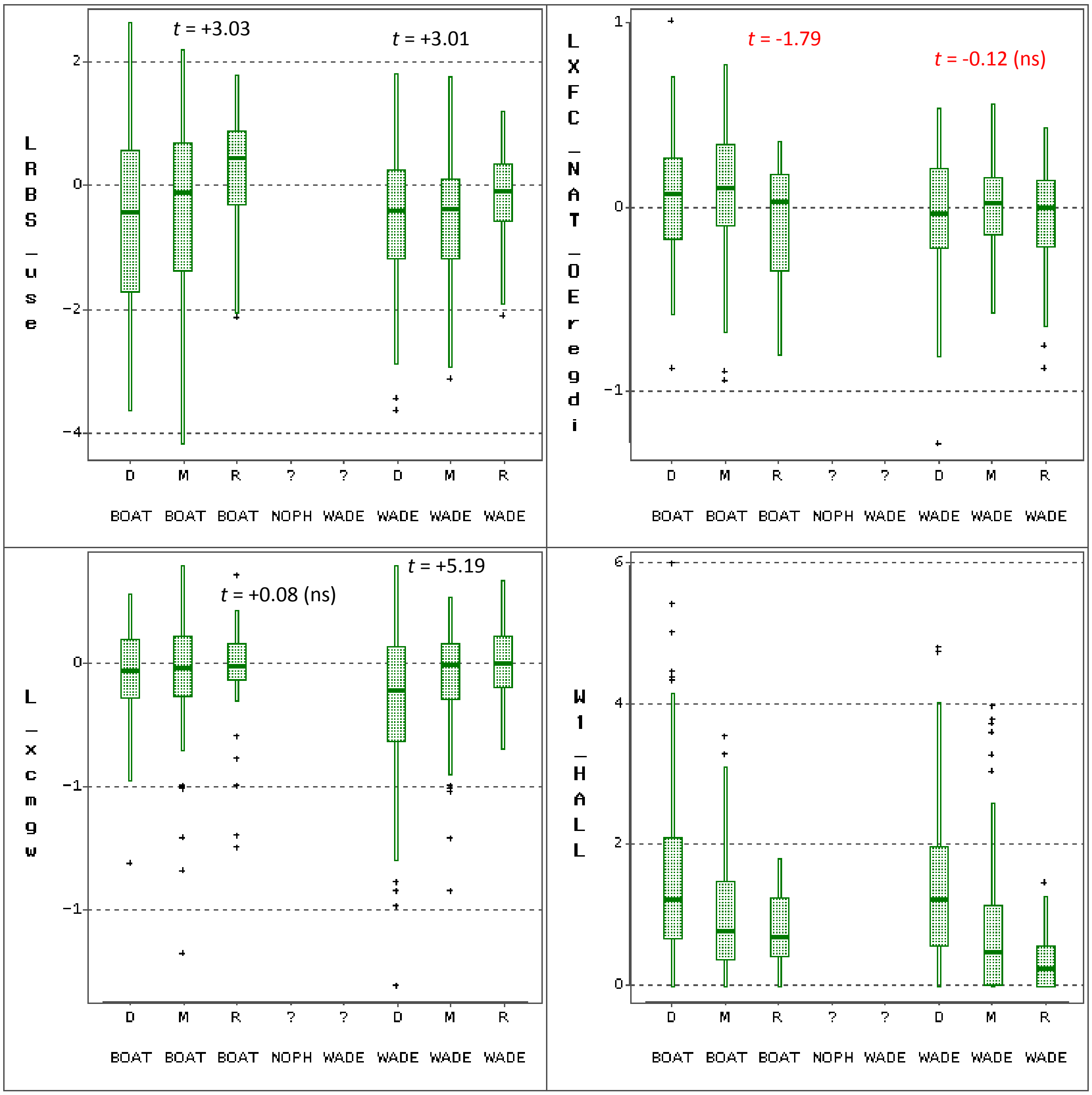


ECO4_pk = CPL

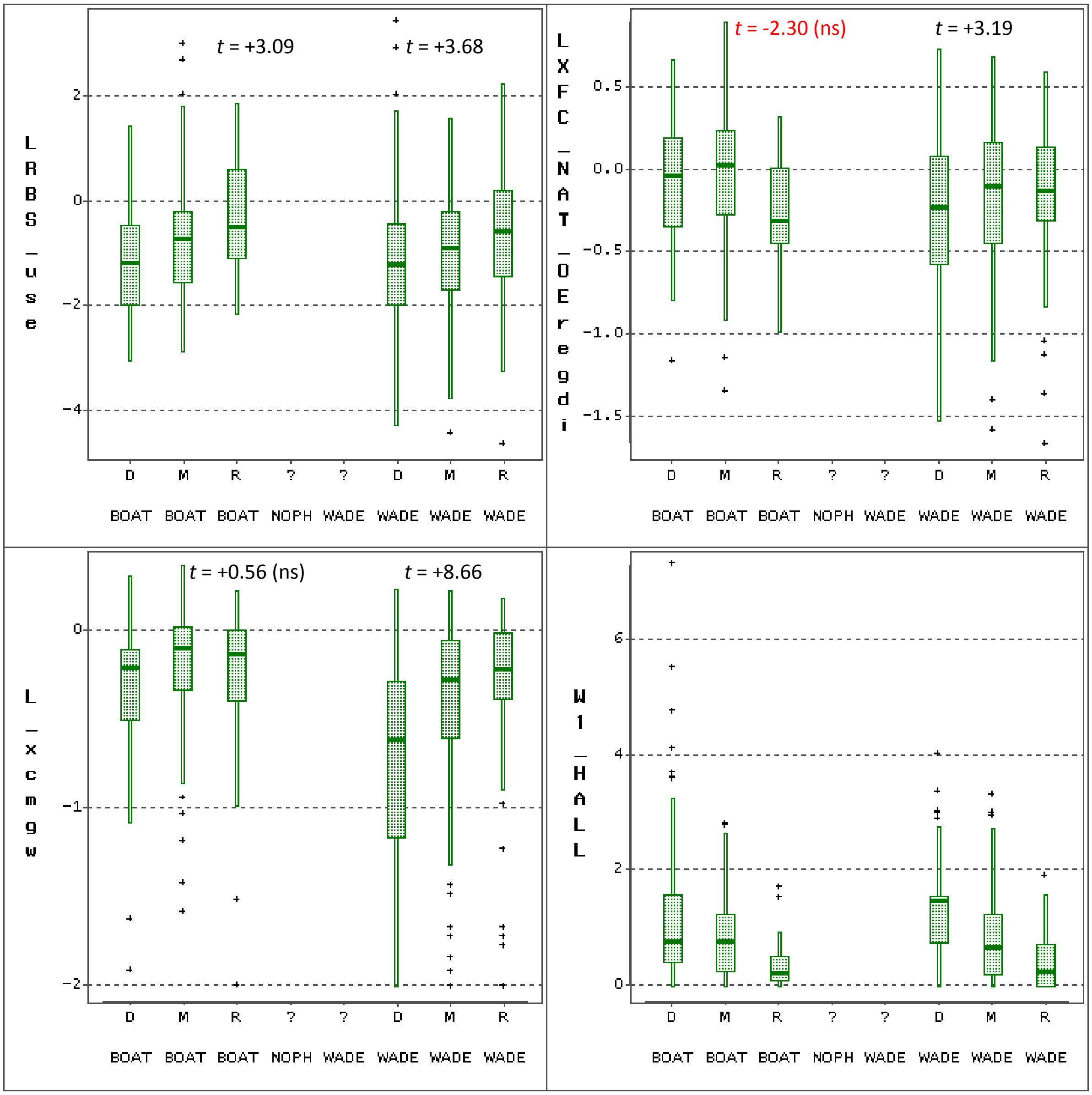


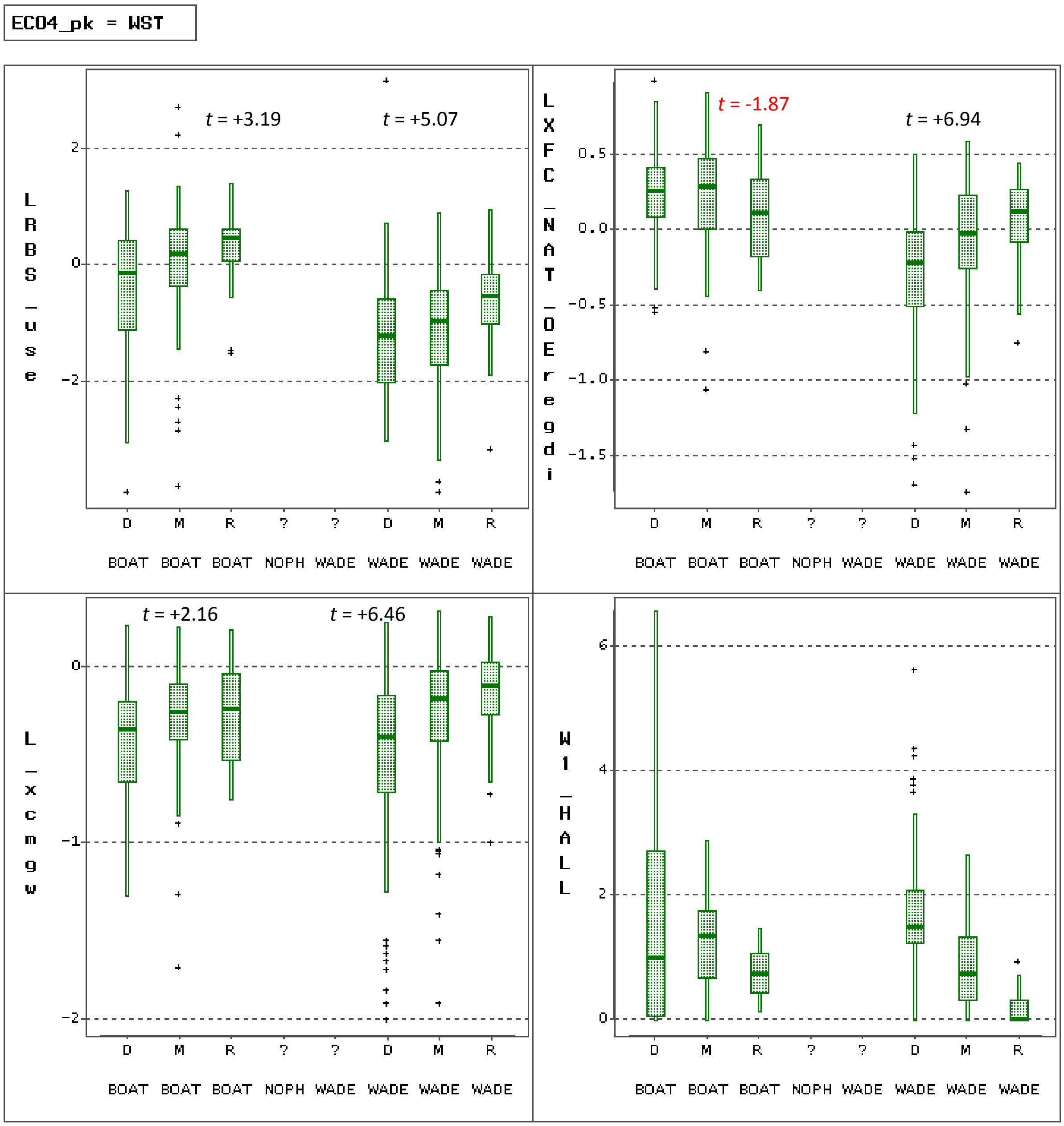
LRBS_use
t = +0.37 (ns)
t =
2
0
-2
LXFC_NAT_OE regdi
t = +0.18
t = +3.78
0.5
0.0
-0.5
-1.0
-1.5
L_xcmgw
t = +4.44
t = +4.71
0
-1
-2
W1_HALL
4
2
0
D M R ? ? D M R
BOAT BOAT BOAT NOPH WADE WADE WADE WADE

ECO4_pk = EHIGH



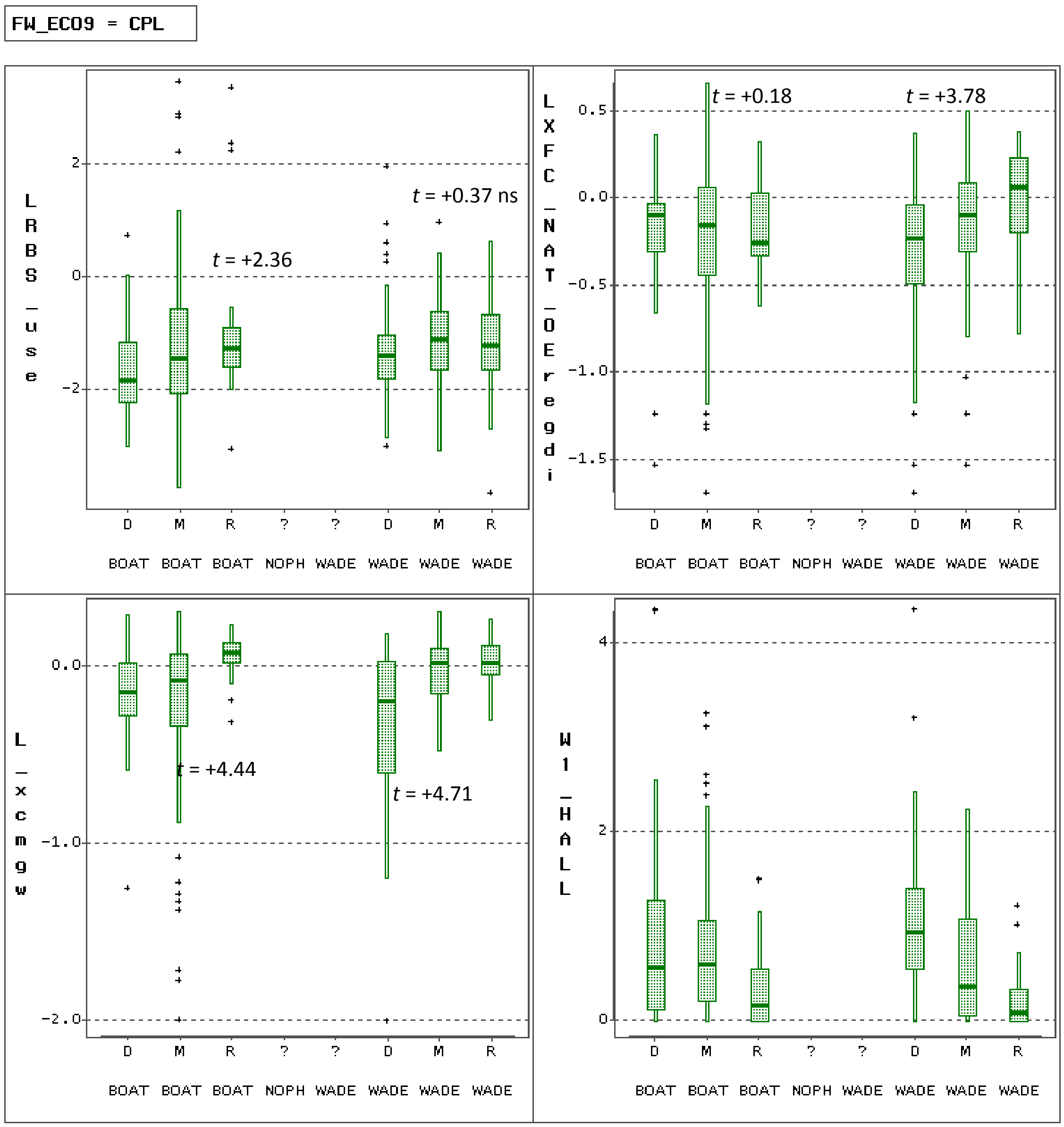
ECO4_pk = INTPLNUMW





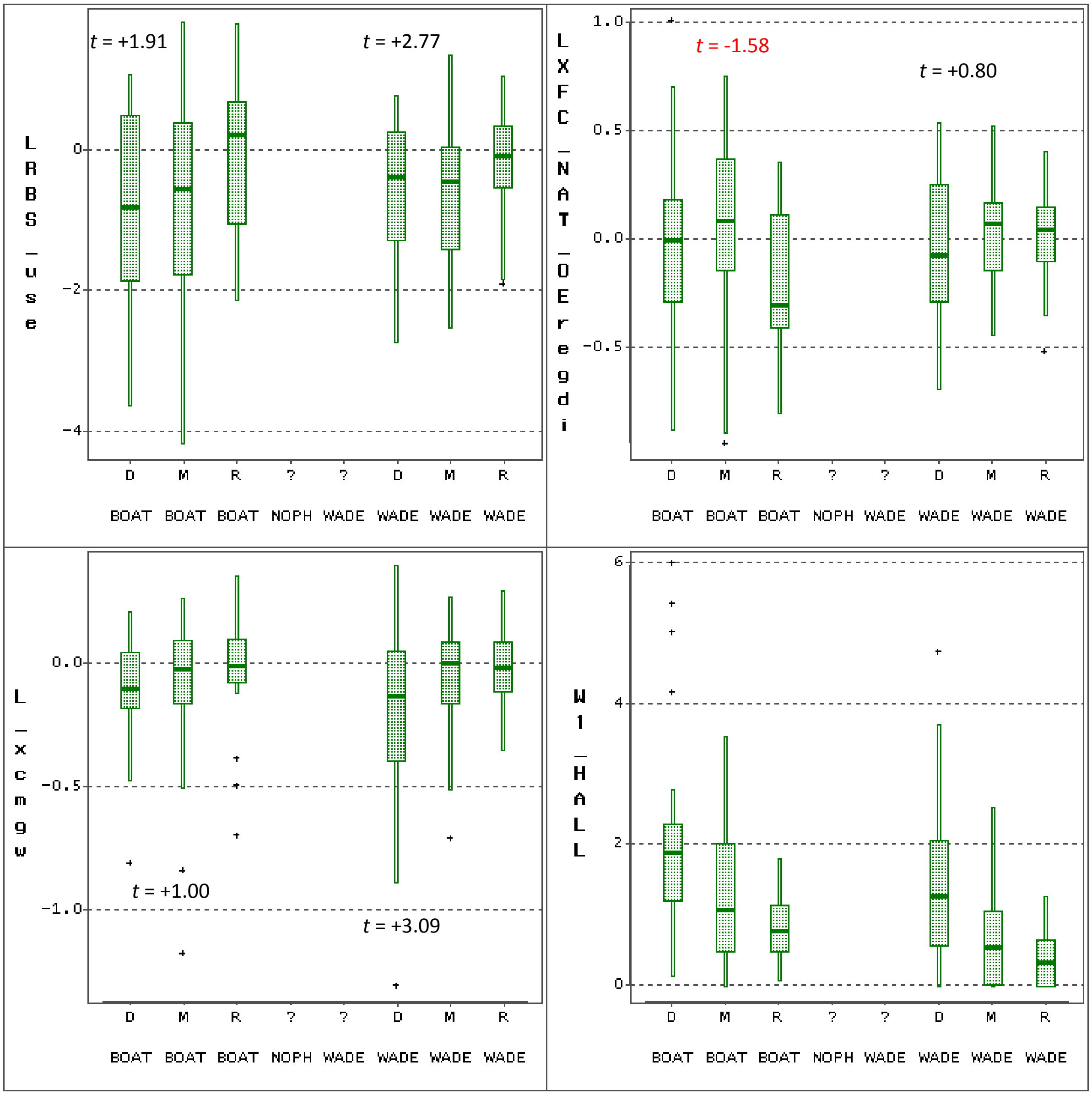
EC04_pk = WST
t = +3.19
t = +5.07
t = -1.87
t = +6.94
t = +2.16
t = +6.46
LRBS_use
LXFC_NAT_OEregdi
L_xcmgw
W1_HALL
D M R ? ? D M R
BOAT BOAT BOAT NOPH WADE WADE WADE WADE

Figures G-2 A through I. Indicator response figures nationally and by ecoregion (FW_ECO9). LRBS_use = LRBS_g08, the indicator of relative bed stability and excess streambed fine sediments. LXCF_Nat_OEregdis = Log10 of Observed/Expected XFC_NAT, the indicator of instream habitat cover complexity. L_xcmgw = Log10(XCMGW), the indicator of riparian vegetation cover and structure. W1_HALL is the proximity weighted indicator of riparian and near-shore human disturbance intensity. Plots show *t* values for the differences between means for reference (R) and disturbed (D) sites. Values shown in red have a sign contrary to expectations.

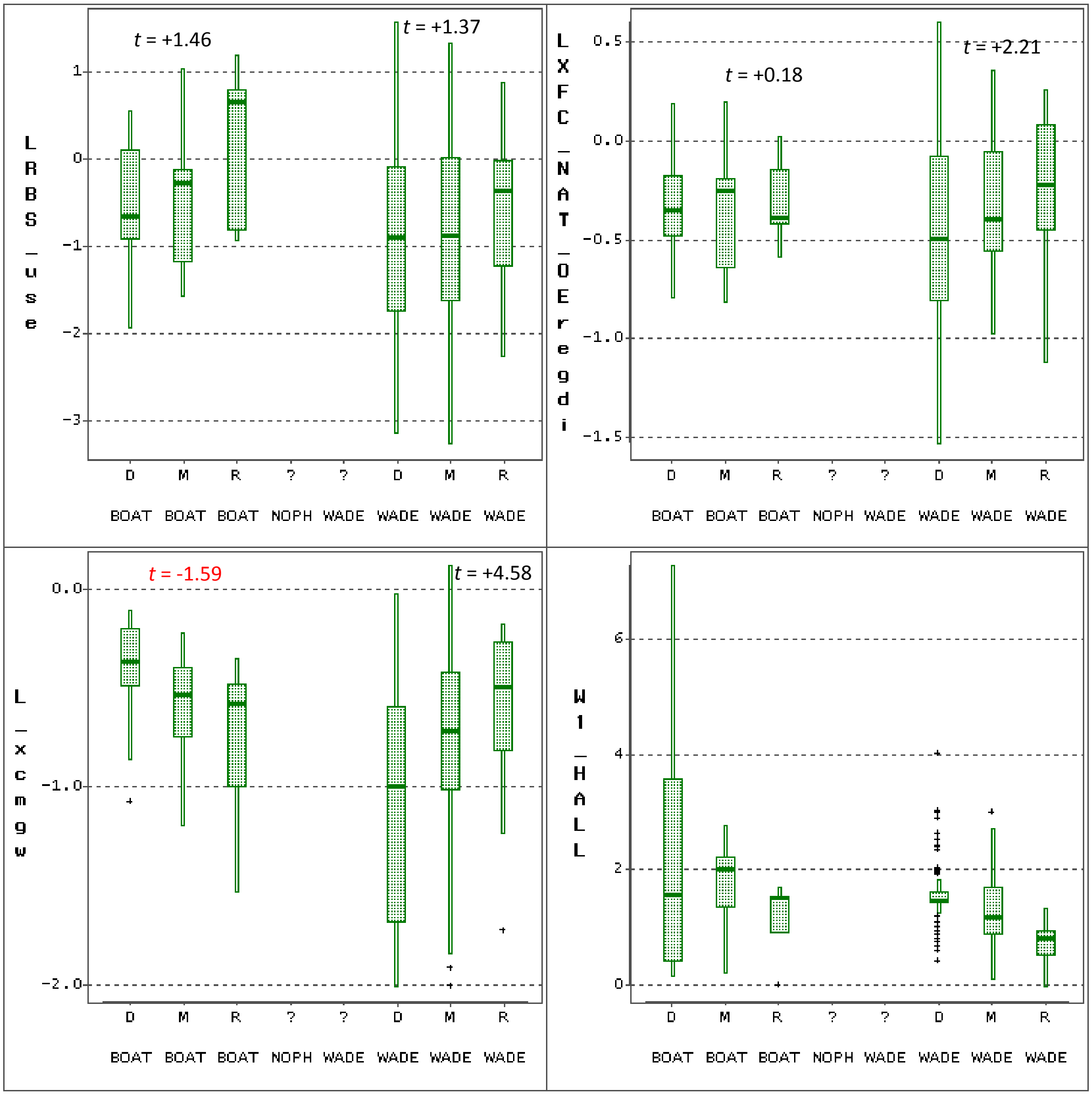


o

FW_ECO9 = NAP

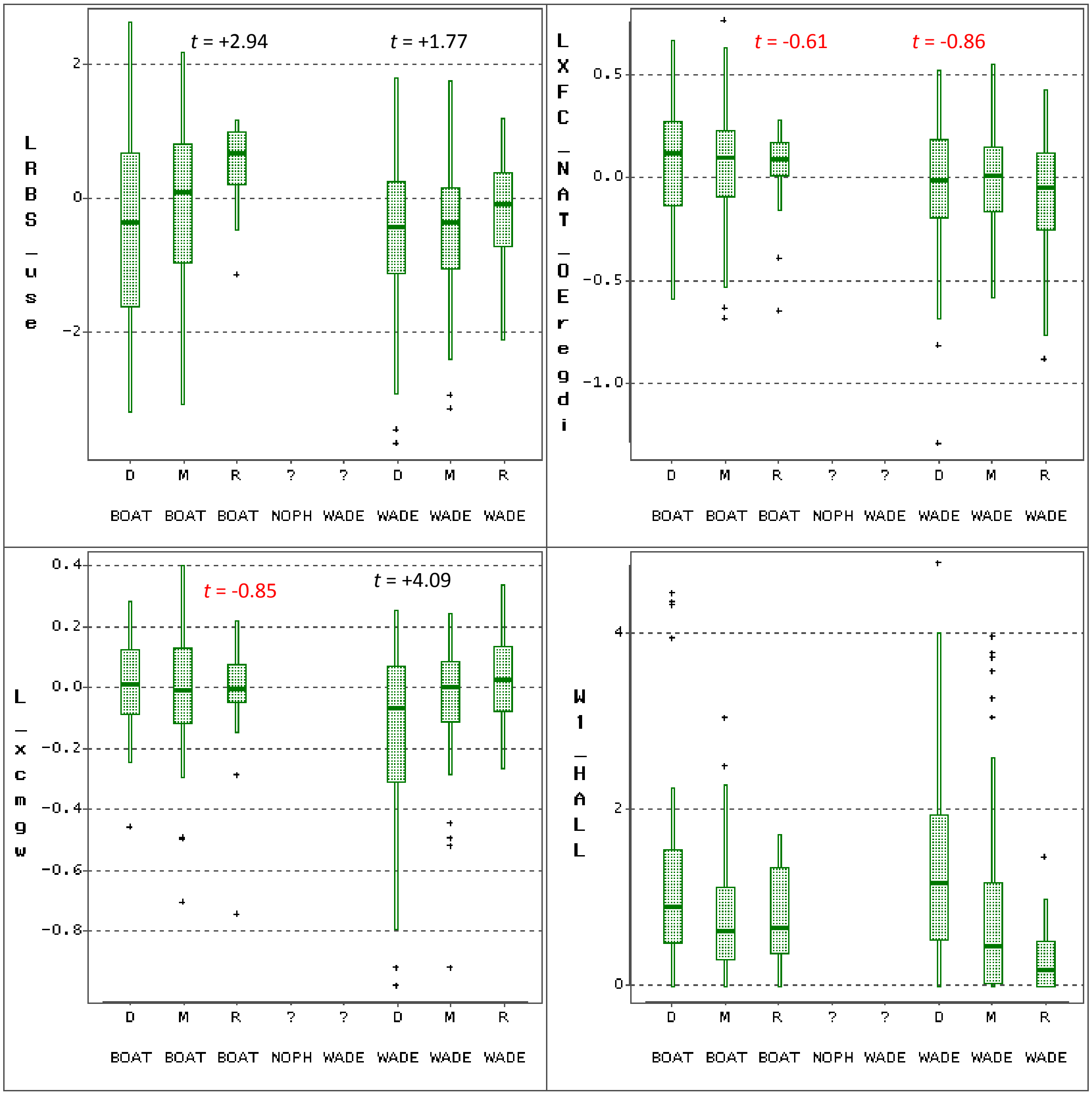

t = +1.91
t = +2.77
LRBS_use
t = -1.58
t = +0.80
LXFC_NAT_OEregdi
t = +1.00
t = +3.09
L_xcmgw
W1_HALL
BOAT BOAT BOAT NOPH WADE WADE WADE WADE
D M R ? ? D M R

FW_ECO9 = NPL



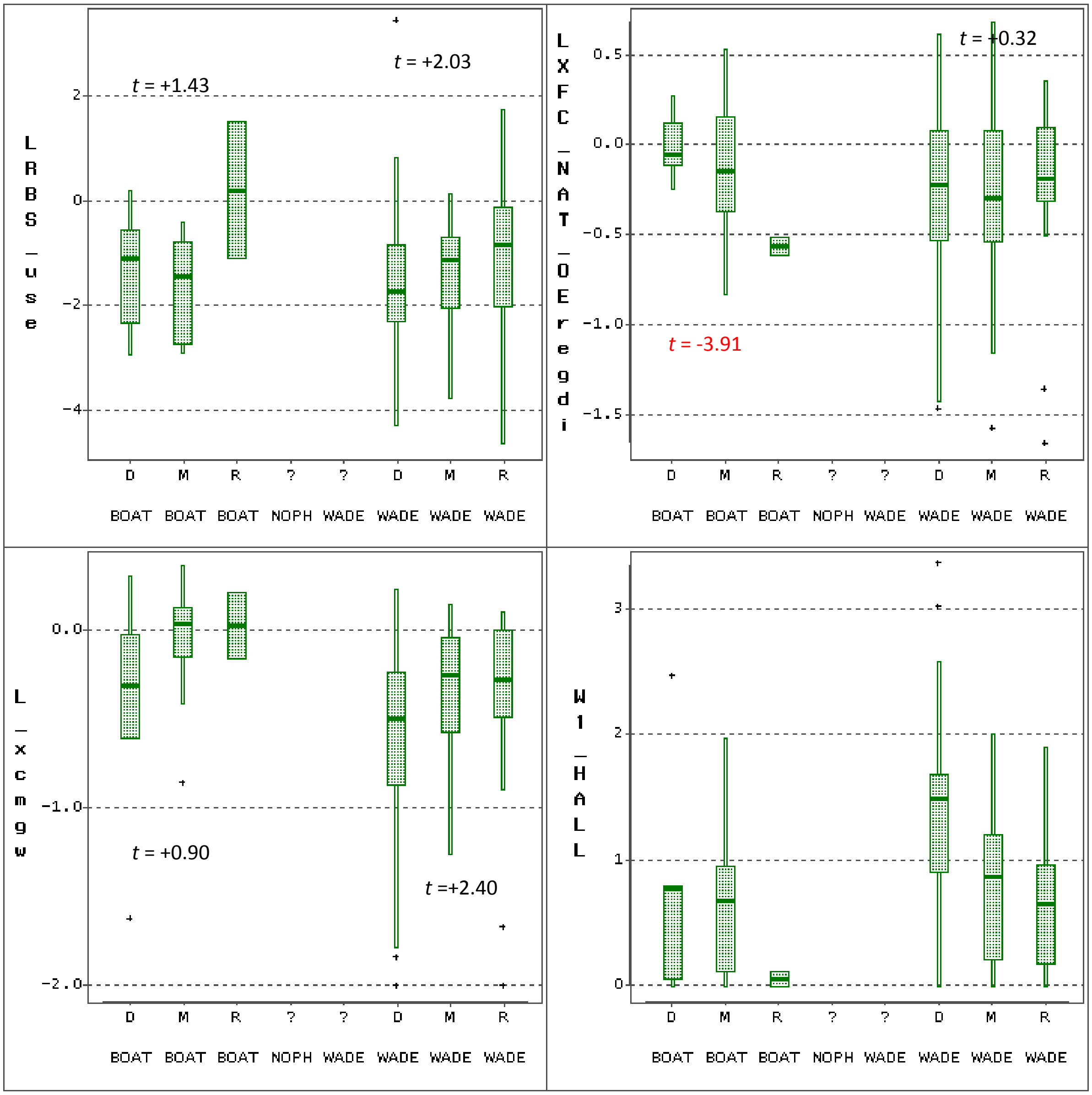
t = +1.46
t = +1.37
t = +0.18
t = +2.21
t = -1.59
t = +4.58
LRBS_use
LXFC_NAT_OEregdi
L_xcmgw
W1_HALL
D M R ? ? D M R
BOAT BOAT BOAT NOPH WADE WADE WADE WADE

FW_ECO9 = SAP



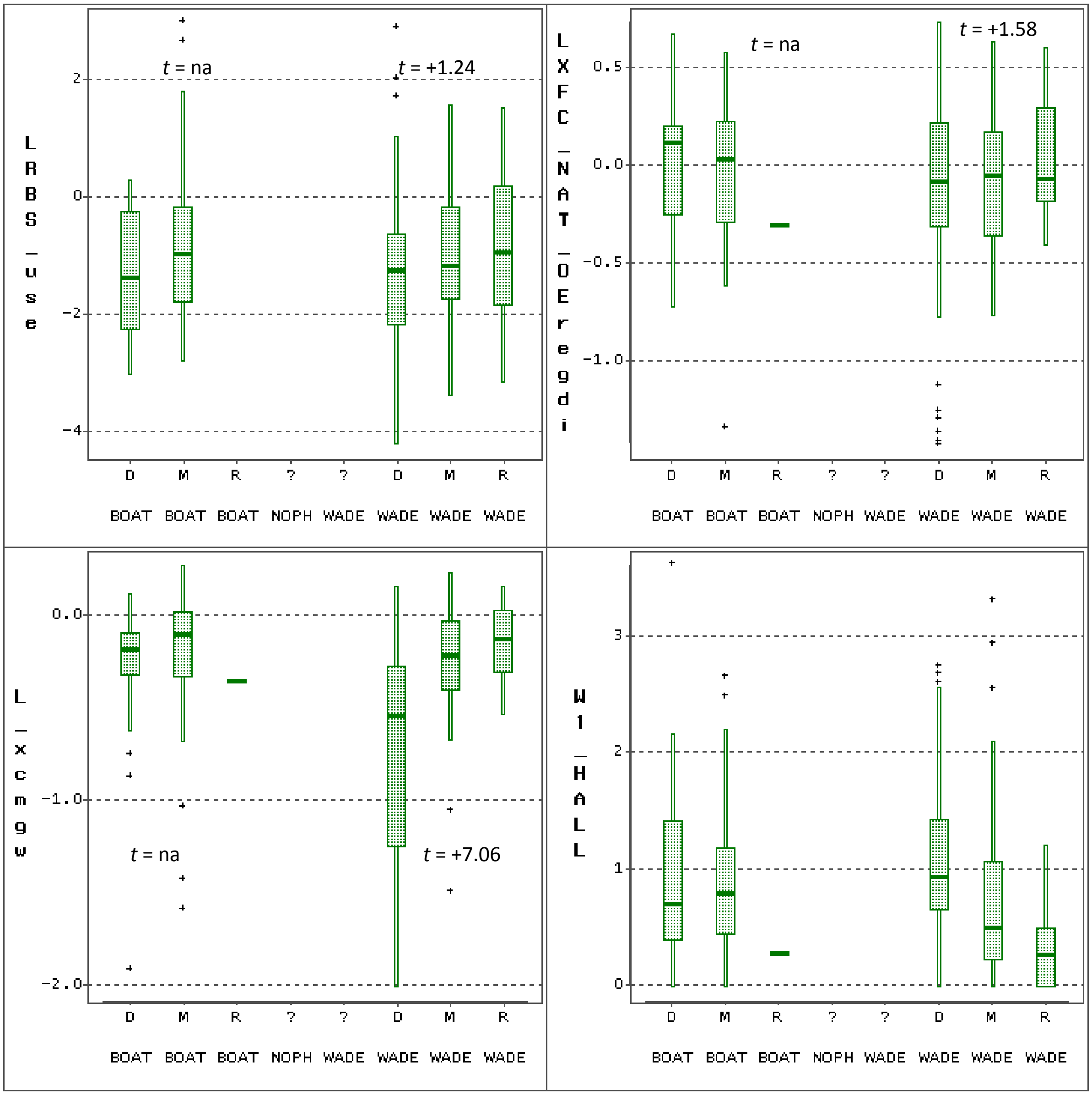
t = +2.94
t = +1.77
LRBS_use
t = -0.61
t = -0.86
LXFC_NAT_OEregdi
t = -0.85
t = +4.09
L_xcmgw
W1_HALL
D M R ? ? D M R
BOAT BOAT BOAT NOPH WADE WADE WADE WADE

FW_ECO9 = SPL



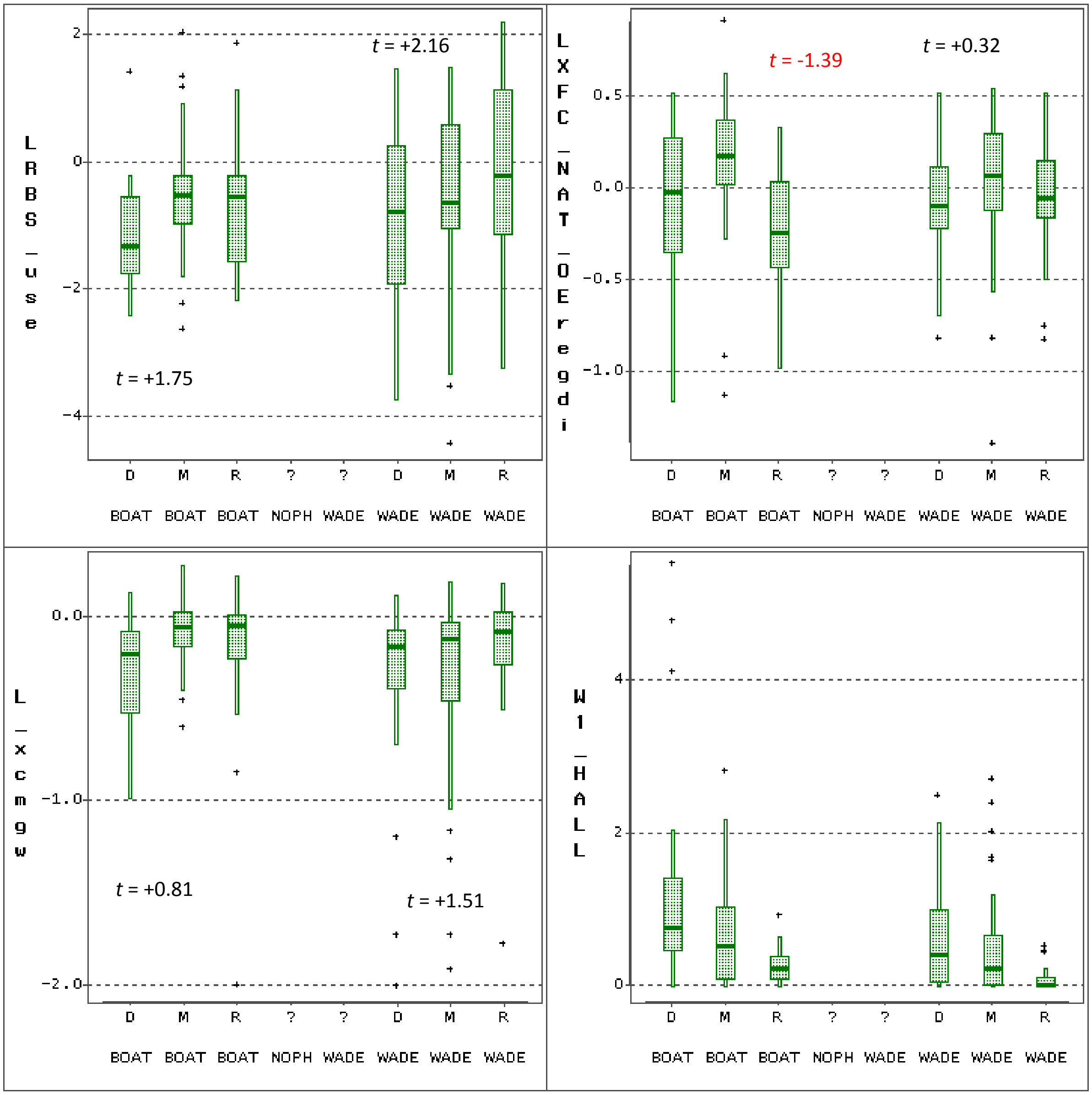
LRBS_use
LXFC_NAT_OEregdi
L_xcmgw
W1_HALL
t = +1.43
t = +2.03
t = +0.32
t = -3.91
t = +0.90
t =+2.40
D M R ? ? D M R
BOAT BOAT BOAT NOPH WADE WADE WADE WADE

FW_ECO9 = TPL



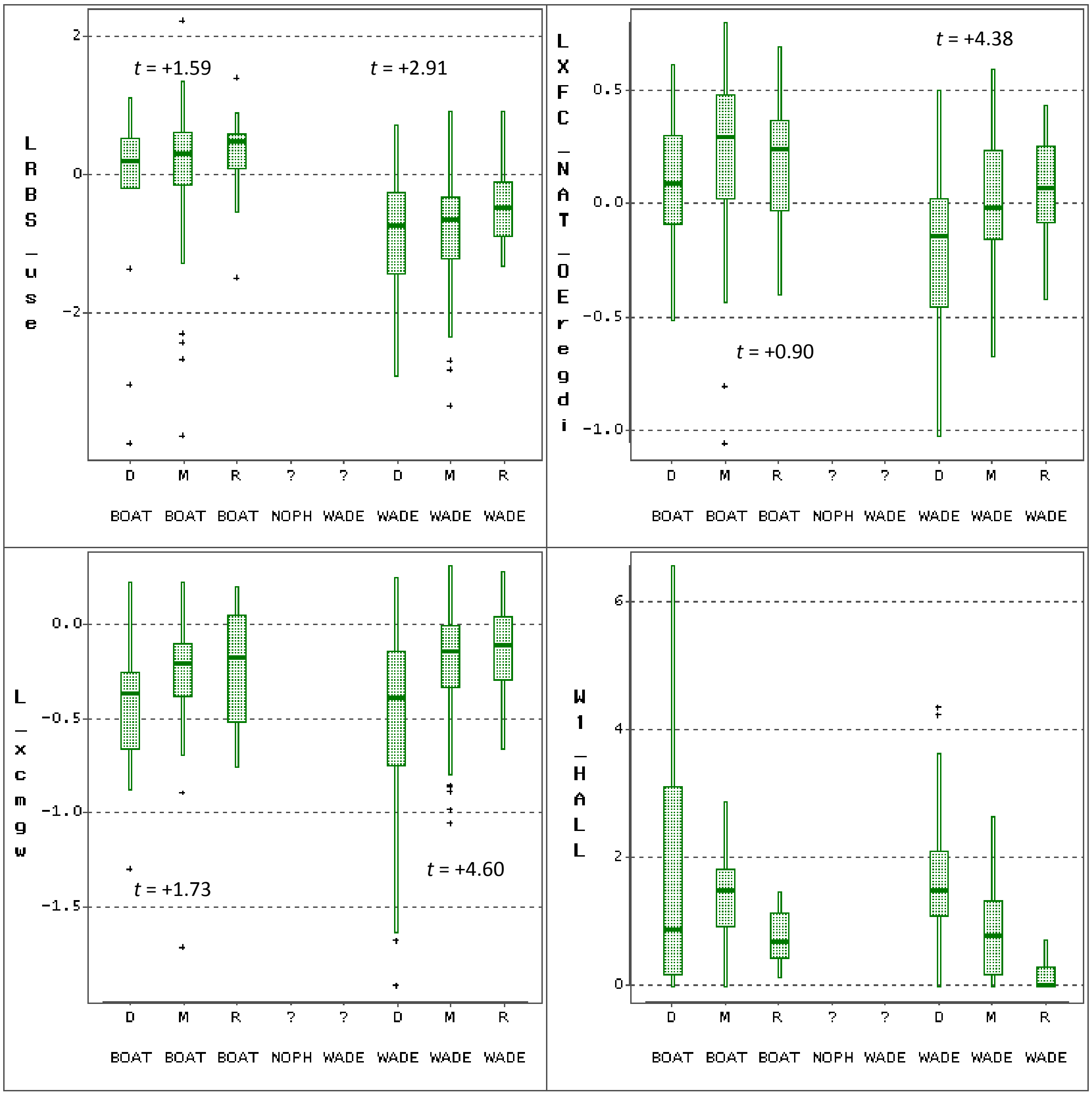
L RBS _use
2
0
-2
-4
t = na
t = +1.24
LXFC _NAT _OE regdi
0.5
0.0
-0.5
-1.0
t = na
t = +1.58
L _xcmgw
0.0
-1.0
-2.0
t = na
t = +7.06
W1 _HALL
3
2
1
0
D M R ? ? D M R
BOAT BOAT BOAT NOPH WADE WADE WADE WADE

FW_ECO9 = UMW

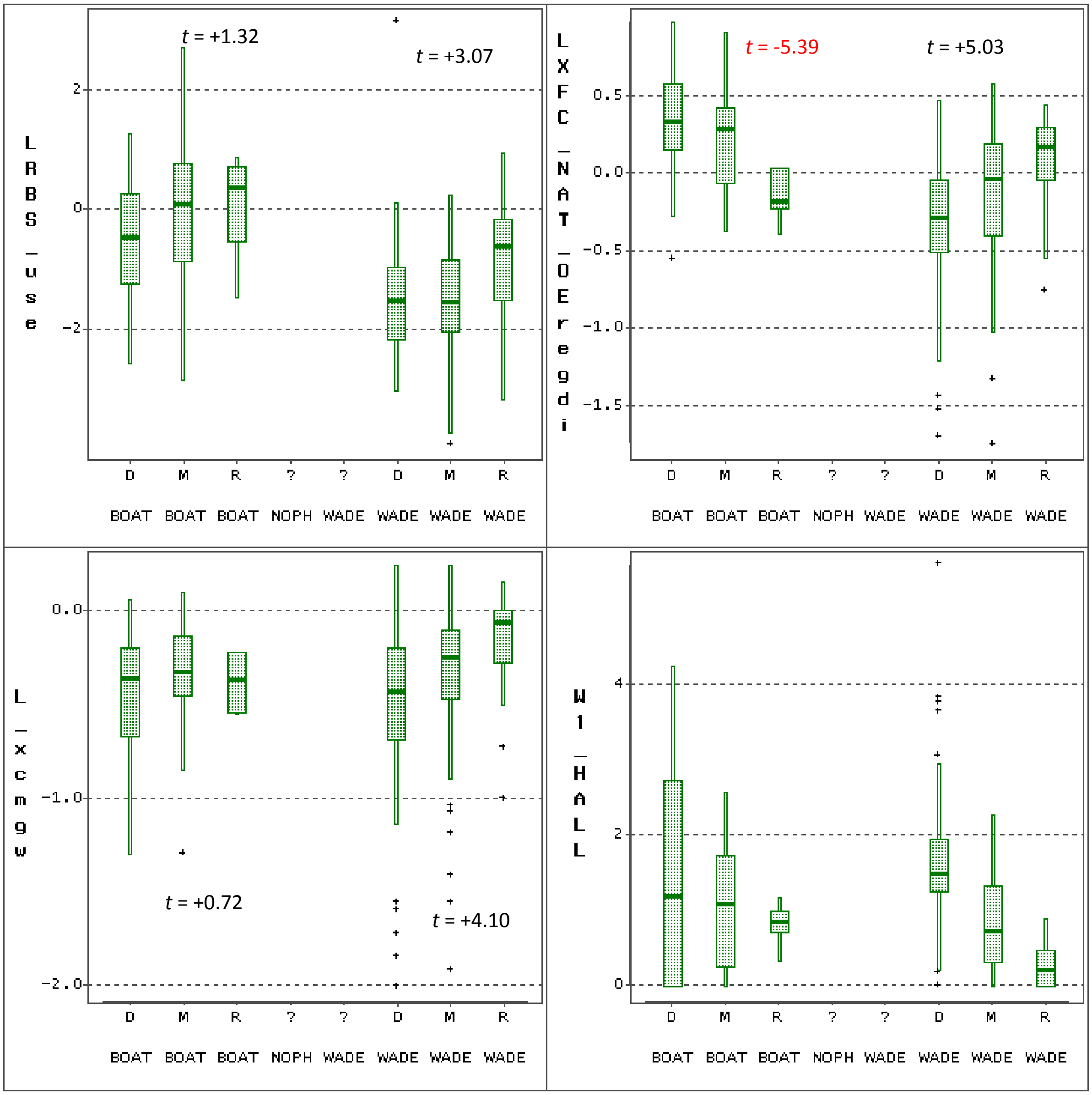


LRBS_use
t = +2.16
t = +1.75
LXFC_NAT_OEregdi
t = -1.39
t = +0.32
L_xcmgw
t = +0.81
t = +1.51
W1_HALL
D M R ? ? D M R
BOAT BOAT BOAT NOPH WADE WADE WADE WADE

FW_ECO9 = WMT


L_RBS_use
t = +1.59
t = +2.91
L_XFC_NAT_OEregdi
t = +4.38
t = +0.90
L_xcmgw
t = +1.73
t = +4.60
W1_HALL
D M R ? ? D M R
BOAT BOAT BOAT NOPH WADE WADE WADE WADE

FW_ECO9 = XER



t = +1.32
t = +3.07
LRBS_use
t = -5.39
t = +5.03
LXFC_NAT_OEregdi
L_xcmgw
t = +0.72
t = +4.10
W1_HALL
D M R ? ? D M R
BOAT BOAT BOAT NOPH WADE WADE WADE WADE

Enterococci Indicator for the NRSA 2008–2009

8.1 Background information

The EPA has developed and validated a molecular testing method called quantitative polymerase chain reaction (qPCR) as a rapid analytical technique for the detection of enterococci in recreational water. NRSA used this method to assess the presence and quantity of fecal indicators in the nation's rivers and streams. EPA then applied the draft threshold and recreational criteria to the enterococci data to assess the recreational condition of streams and rivers.

8.2 Methods

To collect enterococci samples, crews took a water sample for the fecal indicator at the last transect after all other sampling was completed. Using a pre-sterilized 250 mL bottle, they collected the sample approximately 1 m off the bank at about 0.3 m (12 inches) below the water. Following collection, crews placed the sample in a cooler and kept it on ice prior to filtration of four 50 mL volumes. Samples were all filtered and frozen on dry ice within 6 hours of collection. In addition to collecting the sample, crews looked for signs of disturbance throughout the reach that would contribute to the presence of fecal contamination to the waterbody.

This collection and the lab method followed EPA's Enterococcus qPCR method A. Method A describes a quantitative polymerase chain reaction (qPCR) procedure for the detection of DNA from enterococci bacteria in ambient water matrices based on the amplification and detection of a specific region of the large subunit ribosomal RNA gene (lsrRNA, 23S rRNA) from these organisms. Method A uses an arithmetic formula, the comparative cycle threshold (CT) method, to calculate the ratio of enterococcus lsrRNA gene target sequences (target sequences) recovered in total DNA extracts from water samples relative to those in similarly prepared extracts of calibrator samples containing a known quantity of enterococcus cells. Mean estimates of the absolute quantities of target sequences in the calibrator sample extracts are then used to determine the absolute quantities of target sequences in the water samples. CT values for sample processing control (SPC) sequences added in equal quantities to both the water filtrate and calibrator samples before DNA extraction are used to normalize results for potential differences in DNA recovery or to signal inhibition or fluorescence quenching of the PCR analysis caused by a sample matrix component or possible technical error.

8.3 Thresholds

8.3.1 *EPA Thresholds*

To analysis the NRSA data for enterococci using the QPCR method, NRSA is applying the draft EPA thresholds that have been defined and outlined in the EPA report *Recreational Water Quality Criteria* (EPA-HQ-OW-2011-0466).The document contains the EPA's draft ambient water quality criteria recommendations for protecting human health in marine and fresh waters. Please refer to the EPA recommendation document for details of threshold development and values.

8.3.2 *Calibration for qPCR data*

The NRSA data need to have a standardized implementation and calibration of methods for qPCR data to be applied to the existing EPA thresholds. Comparison of results based on CCE reporting unit assumes that target sequence copies (TSC) per calibrator cell is the same for all calibrator cell preparations. If calibrator TSC/cell values are not similar, test sample CCE estimates are not comparable without an adjustment. TSC/calibrator cell values can be determined from standard curves using DNA standards of known TSC concentration. Evidence has been seen in several studies that TSC per cell may vary for different calibrator cell preparations.

Estimates of TSC per calibrator cell for new calibrator cell preparations must also be related back to the values associated with the QPCR criteria values in the pending recreational water quality criteria. Estimates of TSC per calibrator cell generated from standard curves for the National Lakes Assessment and NRSA studies had a mean estimate of 20.41 TSC per cell.

8.4 Literature cited

Cabelli, V.J., A.P. Dufour, M.A. Levin, L.J. McCabe, and P.W. Haberman. 1979. Relationship of Microbial Indicators to Health Effects at Marine Bathing Beaches. *Am. J. Public Health*. 69: 690-696.

Francy, D.S. and Darner, R.A., 1998, Factors affecting Escherichia coli concentrations at Lake Erie public bathing beaches: U.S. Geological Survey, Water-Resources Investigations Report 98-4241, 4 p.

Haugland, R.A., Siefring, S.C., Wymer, L.J., Brenner, K.P. and Dufour, A.P., 2005, Comparison of Enterococcus measurements in freshwater at two recreational beaches by quantitative polymerase chain reaction and membrane filter culture analysis: Water Research, v. 39, p. 559-568.

Oshiro, R.K., Chambers, Y., Pope, M., Miller, K., Grunerud, R., and Keller, K., 2007, Assessment of the effects of holding time on enterococci concentrations in fresh and marine recreational waters and Escherichia coli concentrations in fresh recreational waters: Poster presented at of the American Society for Microbiology Annual Meeting, Toronto, Canada.

Pope, M. L., Bussen, M., Feige, M.A., Shadix, L., Gonder, S., Rodgers, C., Chambers, Y., Pulz, J., Miller, K., Connell, K., and Standridge, J., 2003, Assessment of the effects of holding time and temperature on *Escherichia coli* densities in surface water samples: Applied and Environmental Microbiology, v. 69, no. 10, p. 6201-6207.

U.S. Environmental Protection Agency, 1986, Ambient Water Quality Criteria for Bacteria—1986: Washington, D.C., U.S. EPA 440/5-84-002, 18 p.

U.S. Environmental Protection Agency, 2010. Method A: Enterococci in Water by TaqMan® Quantitative Polymerase Chain Reaction (qPCR) Assay.EPA-821-R-10-004.

Wade, T.J., Calderon, R.L., Sams, E., Beach, M., Brenner, K.P., Williams, A.H., and Dufour, A.P., 2006, Rapidly measured indicators of recreational water quality are predictive of swimming-associated gastrointestinal illness: Environmental Health Perspectives, v. 114, no. 1, p. 24-28.

Human Health Fish Tissue Indicator — Mercury

9.1 Background

Fish are time-integrating indicators of persistent pollutants, and contaminant bioaccumulation in fish tissue has important human and ecological health implications. Contaminants in fish pose risks to human consumers and to piscivorous wildlife. The NRSA fish tissue indicator provides information on the national distribution of selected persistent, bioaccumulative, and toxic (PBT) chemical residues (e.g., mercury and organochlorine pesticides) in predator fish species from rivers 5th order and greater in size of the conterminous United States. For the NRSA report, only the mercury results are presented. For a wide variety of additional chemicals (including selenium, pesticides, PCBs, and other contaminants of emerging concern), analyses are still underway and will be presented in future publications.

The fish tissue indicator procedures were based on EPA's *National Study of Chemical Residues in Lake Fish Tissue* (final report now available) and EPA's *Guidance for Assessing Chemical Contaminant Data for Use in Fish Advisories,* Volume 1 (third edition).

9.2 Field fish collection

The NRSA crews collected fish for the tissue indicator from rivers 5th order and greater in size. Fish tissue samples consisted of a composite of fish (i.e., five individuals of one predator species) from each site. The fish had to be large enough to provide sufficient tissue for analysis (i.e., 500 grams of fillets, collectively). Additional fish criteria for each composite sample included that fish:

- Be of the same species (for each site).
- Satisfy legal requirements of harvestable size (or be of consumable size if there were no harvest limits).
- Be of similar size so that the smallest individual in the composite was no less that 75% of the total length of the largest individual.

Crews were provided with a recommended list of target fish species, though they could choose an appropriate substitute if none of the recommended fish were available (see Table I-1).

Table Ii1. Recommended Target Species for Fish Tissue Collection (in Order of Preference)

	Family name	Common name	Scientific name	Length Guideline (Estimated Minimum)
Predator/Gamefish Species (in order of preference)	*Centrarchidae*	Largemouth bass	Micropterus salmoides	~280 mm
		Smallmouth bass	Micropterus dolomieu	~300 mm
		Black crappie	Pomoxis nigromaculatus	~330 mm
		White crappie	Pomoxis annularis	~330 mm
	Percidae	Walleye/sauger	Sander vitreus/S. canadensis	~380 mm
		Yellow perch	Perca flavescens	~330 mm
	Percichthyidae	White bass	Morone chrysops	~330 mm
	Esocidae	Northern pike	Esox lucius	~430 mm
	Salmonidae	Lake trout	Salvelinus namaycush	~400 mm
		Brown trout	Salmo trutta	~300 mm
		Rainbow trout	Oncorhynchus mykiss	~300 mm
		Brook trout	Salvelinus fontinalis	~330 mm

9.3 Mercury analysis and human health screening values

All fish tissue samples were analyzed for total mercury using a commercially available mercury analyzer that requires only a small amount of tissue (about 1 gram) for analysis. In screening-level studies of fish contamination, EPA guidance recommends monitoring for total mercury rather than methylmercury since most mercury in adult fish is in the toxic form of methylmercury. Applying the conservative assumption that all mercury is present in fish tissue as methylmercury is also more protective of human health. The human health screening value used to interpret mercury concentrations in fillet tissue is 0.3 milligrams (mg) of methylmercury per kilogram (kg) of tissue (wet weight) or 300 parts per billion (ppb), which is EPA's tissue-based water quality criterion for methylmercury. This threshold represents the concentration that, if exceeded, can potentially be harmful to human health. Application of this threshold to the fillet data identifies the number and percentage of river miles in the sampled population for this study that exceed the mercury human health screening value. Results are presented for the miles of 5^{th} order and larger rivers that could not be sampled, and the miles that exceed/do not exceed the human health screening value.

9.4 Literature cited

EPA. 2001. Water Quality Criterion for the Protection of Human Health: Methylmercury. EPA-823-R-01-001. U.S. Environmental Protection Agency, Office of Water, Washington, DC. http://www.epa.gov/waterscience/criteria/methylmercury/document.htmlEPA. G*uidance for Assessing Chemical Contaminant Data for Use in Fish Advisories, Volume 1 (Third Edition)*. November 2000. EPA 823-B-00-007 http://water.epa.gov/scitech/swguidance/fishshellfish/techguidance/risk/upload/2009_04_23_fish_advice_volume1_v1cover.pdf

EPA. National Study of Chemical Residues in Lake Fish Tissue. September 2009. EPA 823-R-09-006. http://water.epa.gov/scitech/swguidance/fishstudies/upload/2009_9_28_fish_study_data_finalreport.pdf

Relative Extent, Relative Risk and Attributable Risk for the NRSA 2008–2009

10.1 Background information

A major goal of the national aquatic surveys being conducted by the Environmental Protection Agency in partnership with the states and tribes is to assess the relative importance, at a regional scale, of stressors that impact aquatic biota. In the NRSA, stressors and their impact to the biological condition were assessed in three different ways. The first looked at the relative extent of the stressor at various geographic scales. The second looked at the relative risk of the stressor to the aquatic biota, and the third assessed the attributable risk of the stressor. The process used to conduct this analysis is described in the attached journal article. This was used for the WSA, the National Lakes Assessment, and the NRSA. While the numbers and results differ across the country, the methods and the processes remained the same for all assessments. Specific questions about the relative, attributable risk or stressors extent should be submitted to the EPA during the comment period. Additional write-ups for the calculations will be provided.

Change Analysis from Wadeable Streams Assessment 2004 to National Rivers and Streams Assessment 2008–2009

11.1 Background information

The 2008/2009 National Rivers and Streams Assessment (NRSA) is the first comprehensive, statistically valid survey of the nation's flowing water resources. It is one in a series of surveys designed to assess the condition of all waters (rivers/streams, lakes, wetlands, and coastal waters). The NRSA is a collaborative effort and partnership between EPA, states, and tribes. Data from this assessment will serve as the baseline for the condition of the nation's rivers and the first change analysis of the nation's wadeable streams.

The sampling design for the NRSA is a probability-based network that provides statistically valid estimates of condition for all rivers and streams with a known confidence. Field crews composed of states, tribes, EPA, USGS, and contractors sampled a total of 2,341 streams and rivers (including reference, base, enhancement and revisit sites) during the summer index period of 2008 and 2009. The survey measures a wide variety of variables intended to characterize the chemical, physical, and biological condition of the nation's flowing waters.

Previously, EPA and partners reported on the condition of all streams in the Wadeable Streams Assessment (WSA). The change analysis examines difference in the population of wadeable streams between the WSA and the NRSA.

11.2 Overall change analysis

Final estimates of differences between the surveys are presented in the percentages of total streams length in the condition categories "good," "fair," and "poor." WSA thresholds were used for each indicator to have a standard threshold for each indicator across surveys. The analysis incorporates a specifically designated set of sampling weights for change to produce regional as well as national estimates.

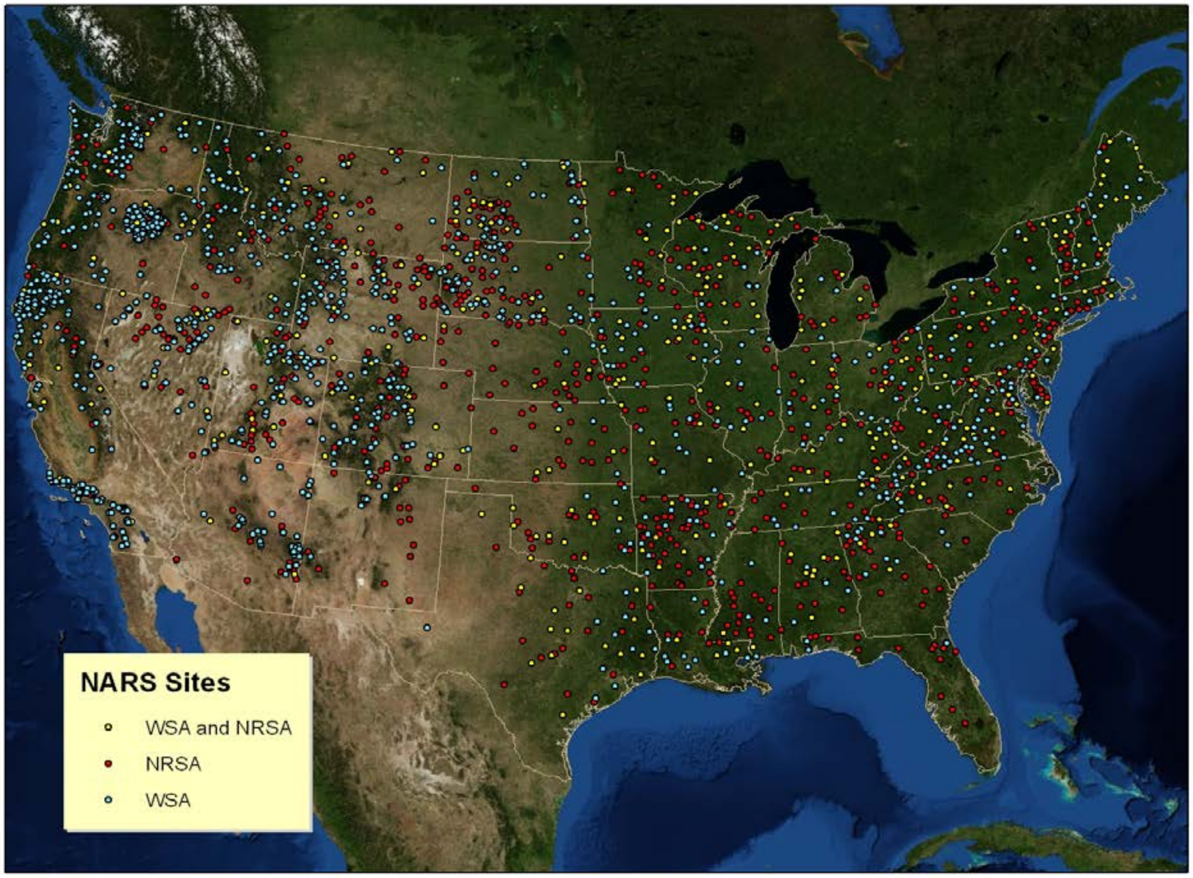


Figure K-1. Sampling locations for the NRSA and WSA. Note the site counts in the legend: substantial sample size (N=2,107). Fewer sites in the west, because fewer streams in the west. Note that 359 of the sites (yellow dots) sampled by the first survey were resampled in the second survey.

Sampling locations in both surveys are shown in Figure K-1. Sites sampled in both surveys have good geographic coverage. Our data analysis is based on one water chemistry sample from each site, from each survey. Samples in both surveys were collected during summer low-flow period, when streams are under maximum stress from high temperature and dewatering. Specific questions on the change analysis should be submitted to EPA through the public comments period.

11.3 Caveats for looking at change analysis

This is the first look at changes in wadeable stream sites across the nation using a statistically valid sampling design. Analysis is still in the initial phases.

- This analysis does not represent a trend; until additional surveys are implemented, we can only look at differences or "changes."

- While this first assessment of the chemistry data shows an increase in total phosphorus, the cause for that increase is still being explored. EPA scientists are examining whether these differences appear to be related to human effects or represent natural variation (flow, etc.). Preliminary analyses have looked at flow issues as well as whether sampling or lab protocols might explain the difference. None of the preliminary work has provided an explanation for the differences.
- These and other possible explanations must still be considered before we draw any conclusion for change from the two surveys.