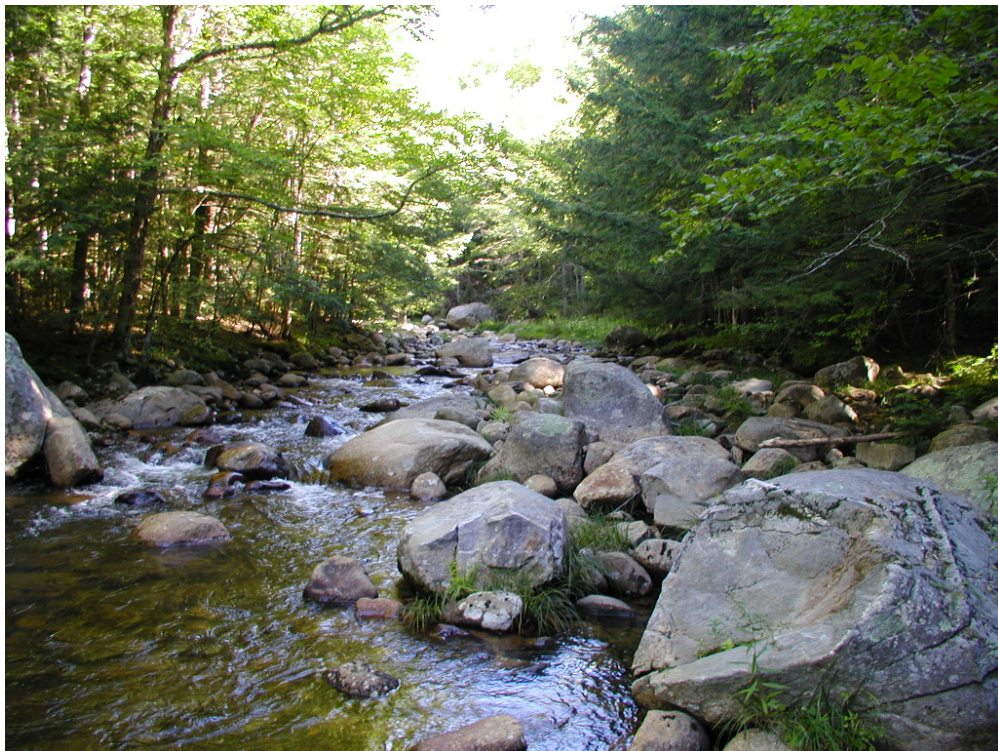


US EPA ARCHIVE DOCUMENT

National Wadeable Stream Assessment: A comparison of assessment outcomes



Outline

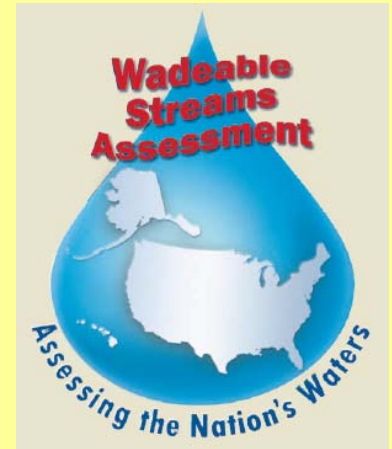
- Overview of WSA results
- Summary of NE participating states sampling & condition tools
- Comparison of assessment outcomes
- Comparison of MMI score for applicable condition tools
- Prediction of MMI scores

THIS IS A WORK IN PROGRESS AND NEEDS YOUR INPUT (THE TRUE SPIRIT OF NEAEB)





Review of the Wadeable Streams Assessment (WSA)



- Conducted in 2004-05
- Probabilistic sampling design including 1,300+ sites nationally
- First national assessment of biological condition using standardized techniques
- Relied on benthic macroinvertebrates as primary biological indicator assemblage
- Included chemical and physical measures as indicators of stream quality and potential stressors to biological communities
- Results presented at national, major climatic regions, and ecological regions

WSA Findings: National

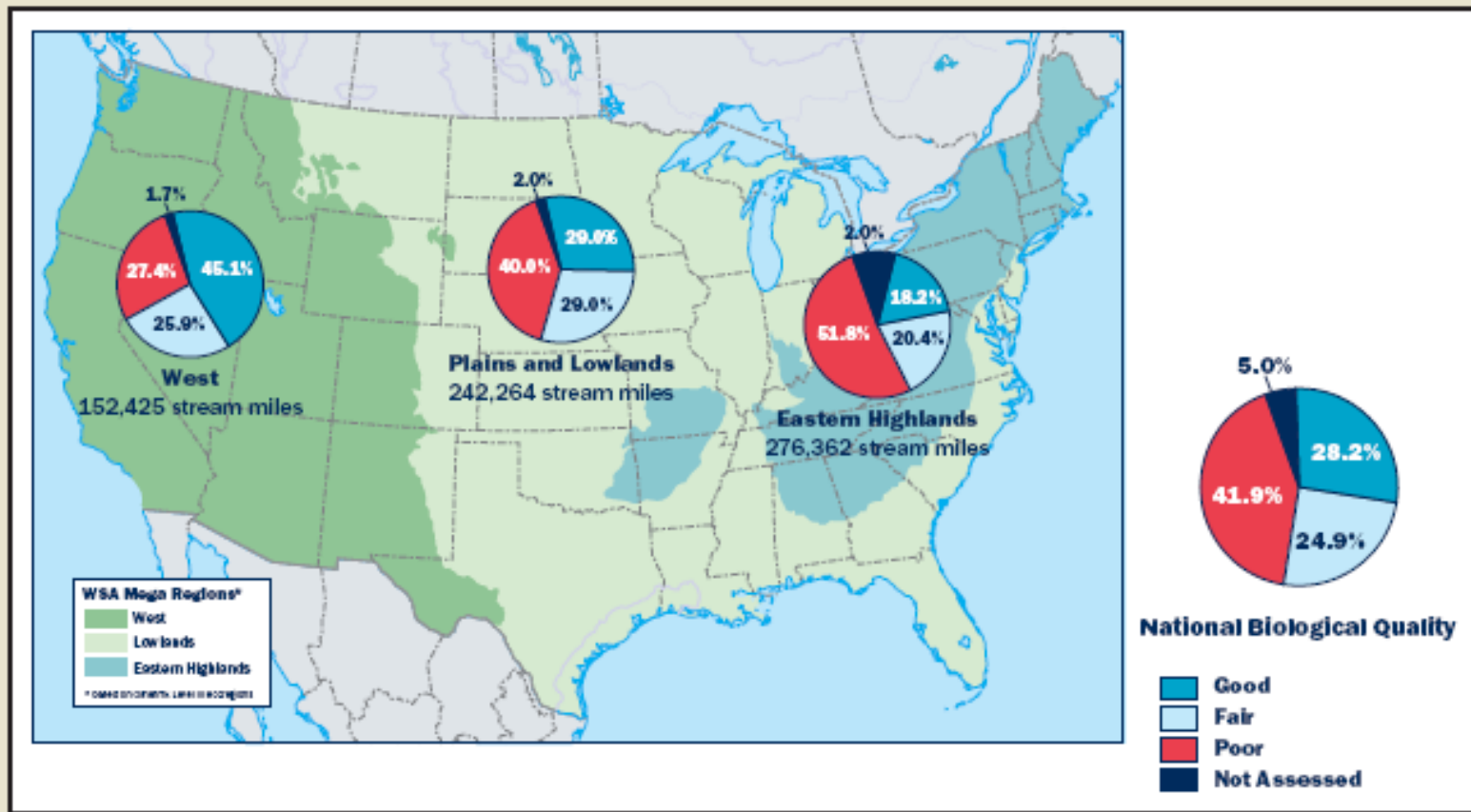


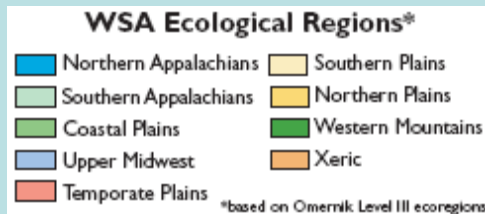
Figure 1. Condition of wadeable streams

- 28% good, 25% fair, 42% poor, 5% unassessed – nationally
- 20% good, 18% fair, 52% poor, 2% unassessed – Eastern Highlands
- Phosphorus, Nitrogen, Excess sedimentation, Riparian disturbance most frequently observed cause of biological community degradation.

WSA Ecoregional Findings for Northern Appalachians (NAP)



- 85 sites
- 97,913 stream miles
- 13% good, 15% fair, 45% poor, 27% unassessed
- 45% high P, 45% high
- 20% high riparian disturbance, 26% poor vegetative cover
- 29% “poor” instream sediments



New England States participation in the WSA

- Cooperative agreement with New England Interstate Water Pollution Commission (NEIWPCA)
- Decision to complete methods comparability data collection at national sites
- CT, ME, NH, and VT
- Macroinvertebrate samples collected at each site using up to 6 methods
- Laboratory sample processing done using each respective entity's method
- Sites within each state sampled by state biologist using up to 4 methods
- Samples sent to contract laboratory for processing according to prescribed procedure

NE States WSA Comparability Study Goal

?

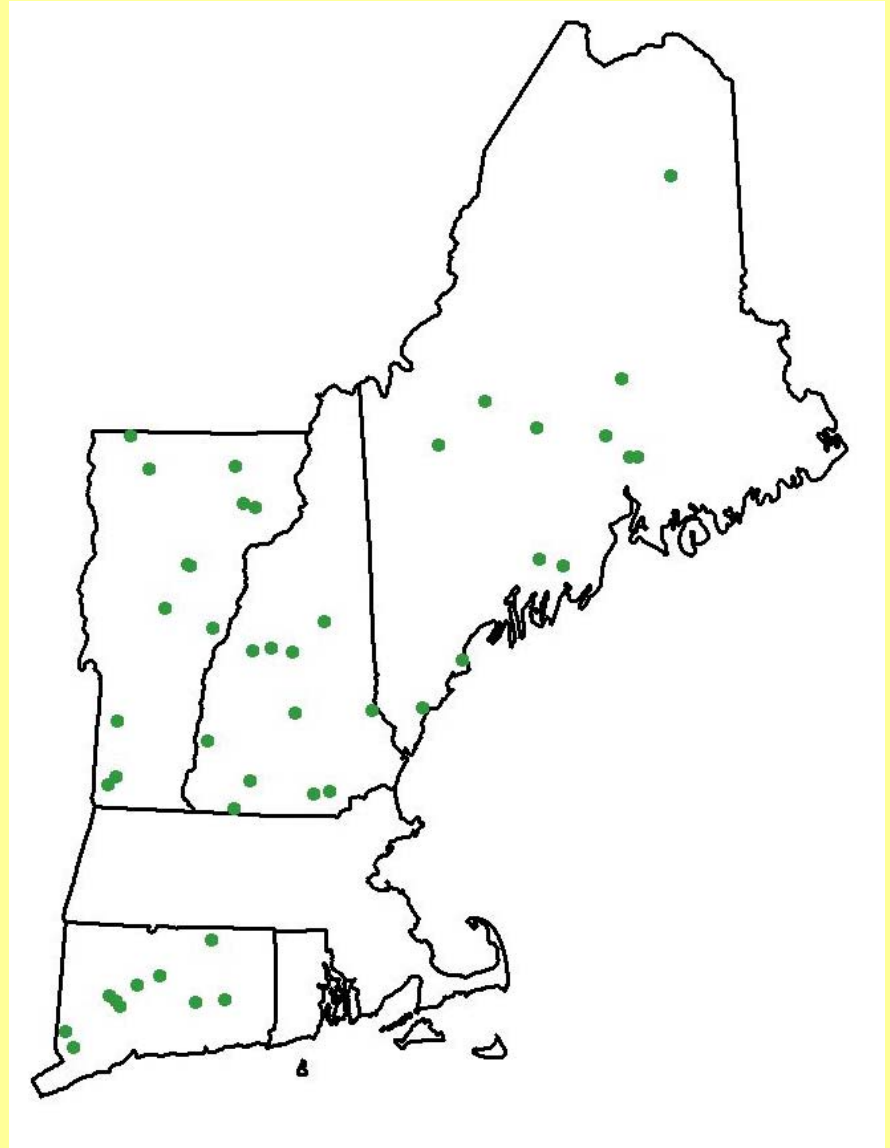
Potential Levels of Comparability:

1. **Taxonomic Composition:** Do the collection and processing techniques result in a sample with a similar “makeup”? Comparison of raw data. See NEWS report appendix I (Jessup and Gerritsen)
2. **Ecological Structure and Function:** Do states measure similar ecological components and are these measures comparable? Compare individual components of state indices and apply Biocondition Gradient (BCG) as common translator (See NEWS report for initial application)
3. **Assessment outcome:** Once samples are processed and ecological attributes computed do states make similar assessment decisions? Check the 303(d) list, but state-specific

Thus, WSA comparability study provided an opportunity to answer the basic question: **Would “x” state list this site?**

New England WSA sample locations and specifics

- 46 sites sampled in CT, ME, NH, and VT
- Sampling completed according to state protocols
- WSA contractors sampled sites using WSA methods
- NEWS data available for limited subset of sites, but not analyzed here
- Subset of sites also considered “reference” condition



Summary of Methods

CT:

- Rectangular net (800µm mesh)
- 12 riffle kicks in fall (Oct.)
- Gridded tray subsampling
- 200 fixed count minimum
- Genus (species) taxonomic endpoint

NH:

- 3 rock baskets
- 6-8 week incubation; 500µm sieve
- Gridded tray subsampling
- ¼ sample minimum; 100 fixed count minimum
- Genus taxonomic endpoint

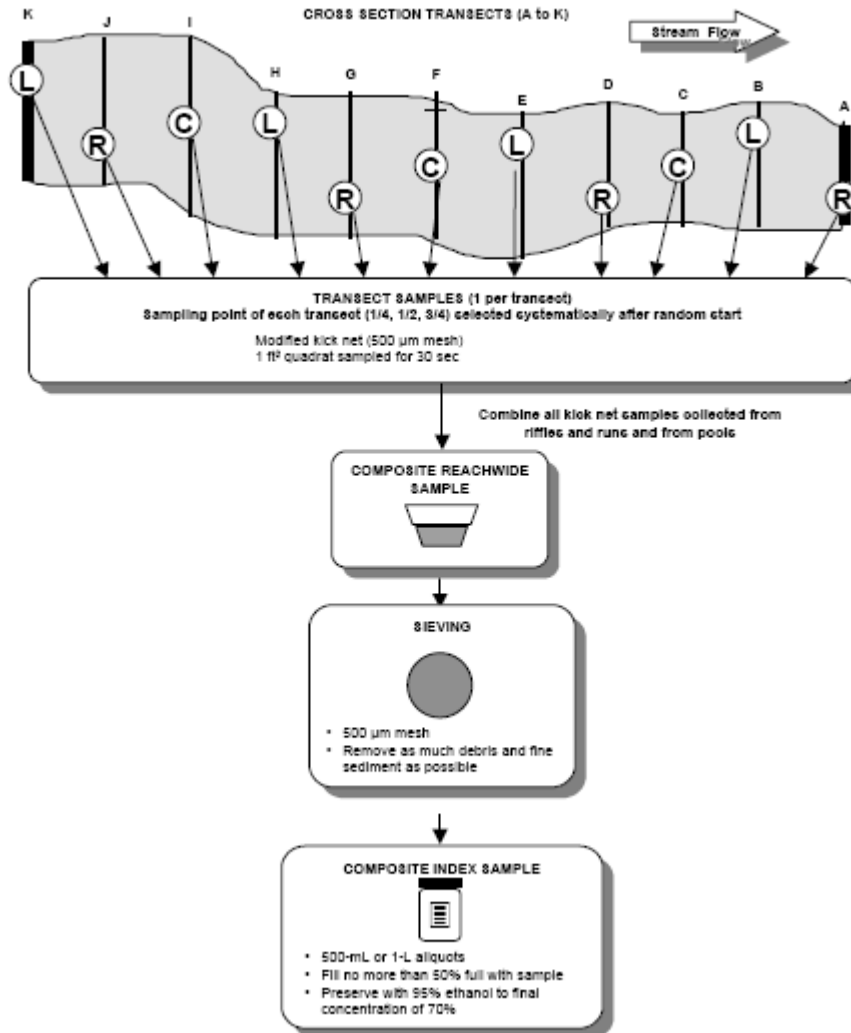
ME:

- 3 rock baskets (bags)
- 28 day incubation
- 600µm mesh sieve for sample preparation
- Entire sample processed
- Genus (species) taxonomic endpoint

VT:

- D-frame net (500µm mesh)
- 4 riffle kicks in fall (Oct.)
- Gridded tray subsampling
- ¼ sample minimum; 300 fixed count minimum
- Genus (species) taxonomic endpoint

WSA Invertebrate Collection and Processing Protocol

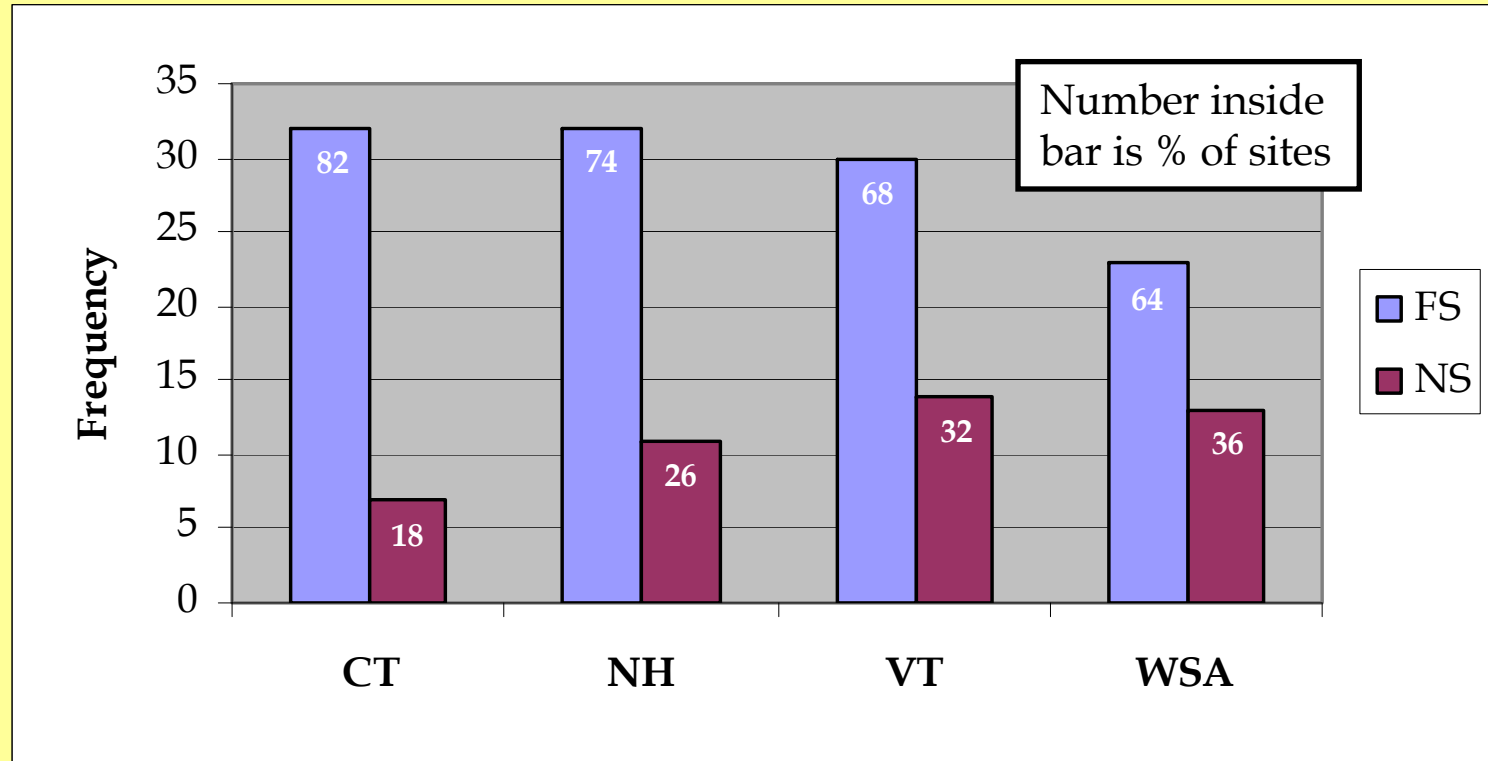


- EMAP
- D-frame net; 500µm mesh
- Multihabitat at discrete points at 11 transects within study reach
- Sample composited into single container
- Gridded sorting
- 500 individual fixed count
- Genus ID end point

Macroinvertebrate Community Condition Evaluation

State	Theshold(s)	Condition "tool"	ALU listing determination
CT	Single w/ grey area requiring BPJ	Multi-metric index; 7 metrics; 0-100; average metric score	Above/Below threshold; threshold based on correspondence with BCG tier rating (Tier 4) + BPJ
ME	Multiple based on TALU system	Multivariate discriminate model; prob. membership to "stream class"	Assignment to correct "stream class" based on model predictions + BPJ
NH	Single w/ consideration for reference condition variability	Multi-metric index; 7 metrics; 0-100; average metric score	Above/Below threshold; threshold set at 25 th percent of reference minus 90% confidence interval + BPJ
VT	Multiple based on TALU system	Multi-metric evaluation; 8 metrics; evaluated independently	Above/Below "stream class" criteria; Based on evaluation of frequency of metric attainment; Individual metric thresholds based on reference condition; + BPJ
WSA	Single based on narrative rating	Multi-metric index; 6 metrics; 0-100; sum of metric scores	Interpretation; Above below "poor" narrative category; Narrative categories established based on reference condition

Overall Comparison of Assessment Outcomes



Assessment Outcomes Do Not Differ
(chi-square: $p > 0.05$)

Levels of Agreement

Number of Assessment Outcomes	Frequency
2	4
3	22
4	20

Status	Complete	Incomplete	Split
FS	26	6	---
NS	7	6	---
Total	32	12	1

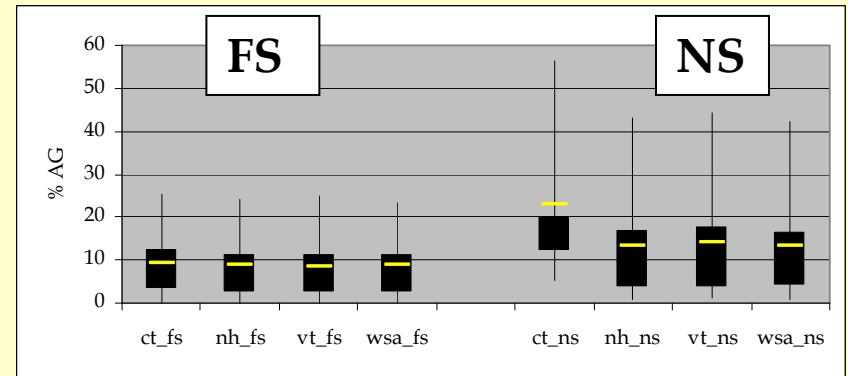
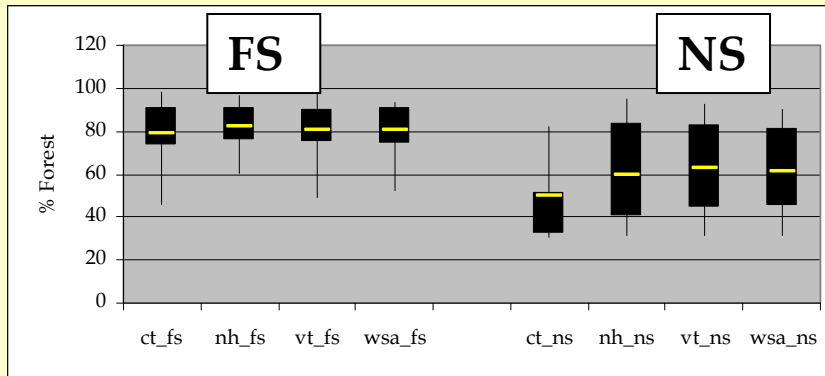
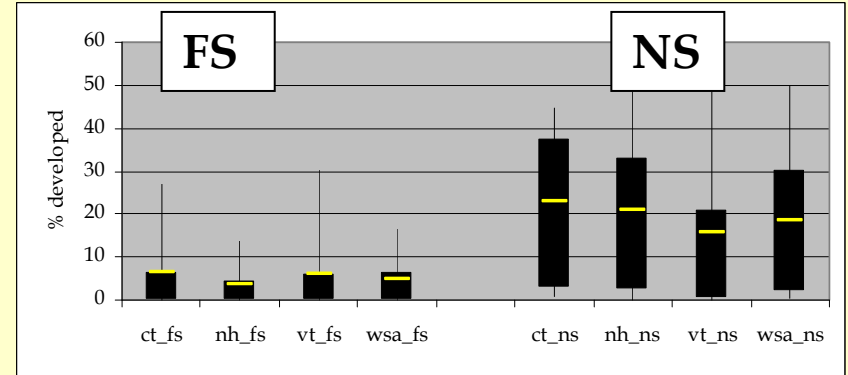
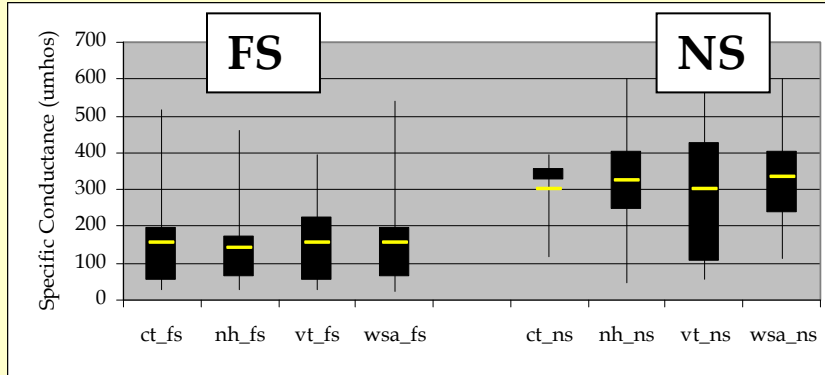
Agreement more common for FS than NS
Chi-square: $p < 0.10$

Dissenting Opinions (Assessments)

Status	CT	NH	VT	WSA
FS	1	4	1	0
NS	0	3	2	1

Dissentions not equally distributed
Chi-square: $p < 0.01$

Assessment Endpoints vs. Environmental Conditions

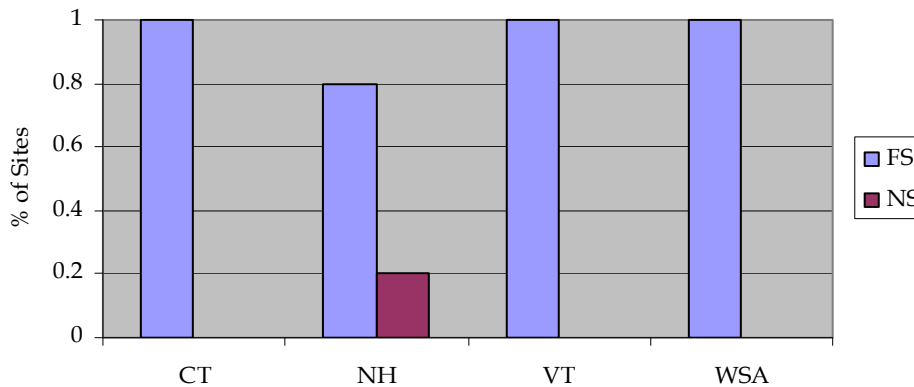


- FS sites coincide w/ lower cond., % develop, %Ag, higher % forest
- Distributions, means similar between states (WSA) for FS and NS sites

Assessment Status: Reference vs. Impacted

- Reference and Impacted rating based on non-biological state assignment
- 21 of 46 sites assigned to Reference (10) or Impacted (11)

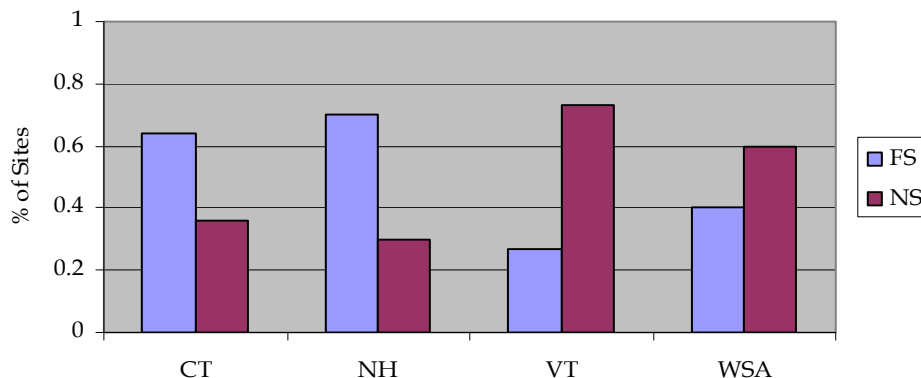
Reference



Regardless of condition index
most sites receive FS rating

Hypothesis: High quality sites
are not sensitive to condition
index (i.e. even a mediocre
index will give FS rating)

Impacted



Impacted sites do not always receive
NS rating & indices vary in status
assignment

Hypothesis: Impacted sites vary in
quality & are more sensitive to
condition index

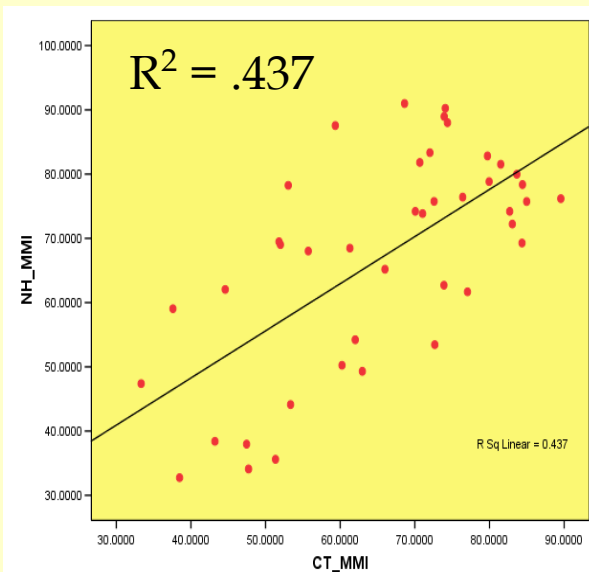
MMI comparisons: CT, NH, WSA

Questions of Interest:

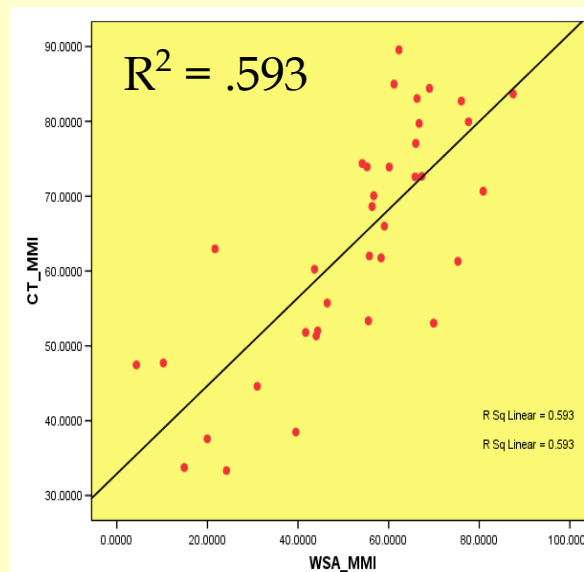
1. Can MMI scores from one index be converted to another?
2. How do converted assessment outcomes compare to actual scores?
3. What is the range of variability around converted scores and how does it affect the assessment outcome?

Conversion of MMI scores: Simple linear regression

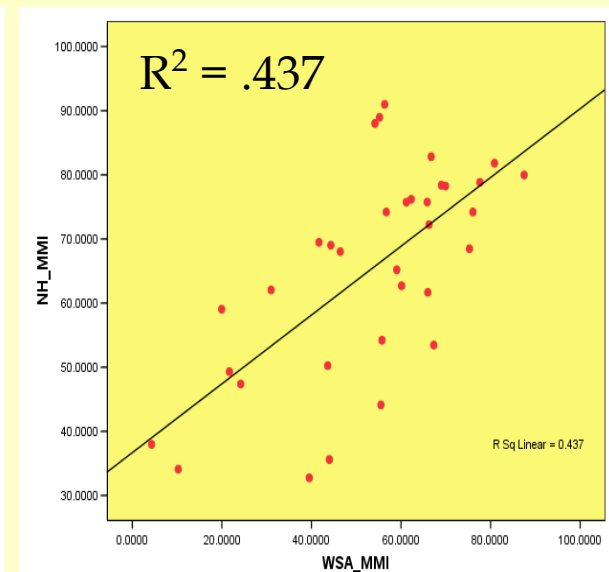
CT/NH



WSA/CT



WSA/NH



Predicted vs. Actual Assessment Outcomes

1. Use regression equation to “convert” one index score to another
2. Compare predicted to actual assessment outcome

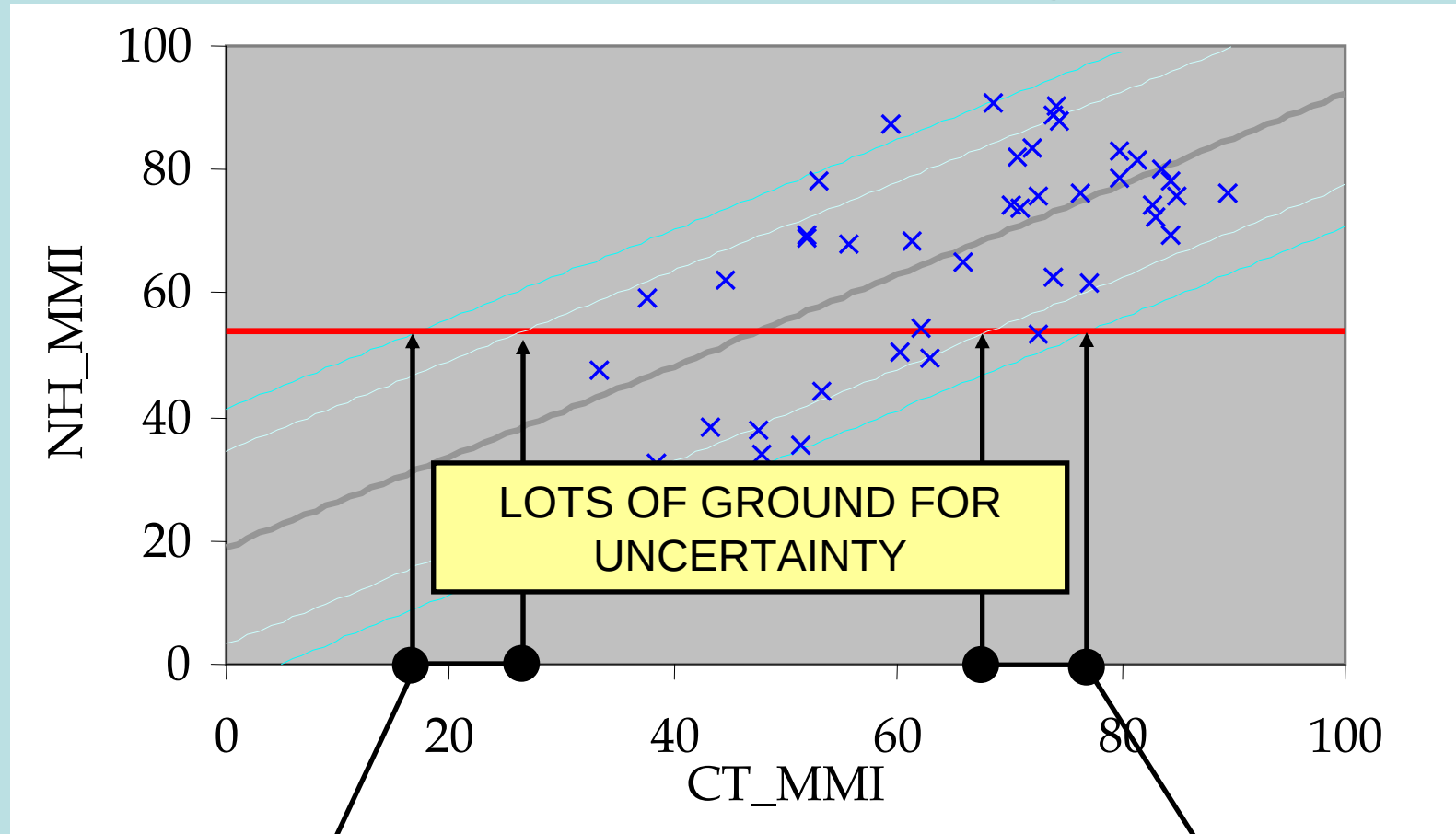
Example:

Site x: CT MMI score = 33.3; predicted NH score = 43.3, actual score = 47.4; NH threshold = 54; both NS status

Results:

	Agree	Disagree	Pred_FS / Act_NS	Pred_NS / Act_FS
CT / NH				
Pred NH vs. Actual	36 (84%)	7 (16%)	3	4
WSA / CT				
Pred CT vs. Actual	25 (69%)	11(31%)	0	1
WSA / NH				
Pred NH vs Actual	28 (82%)	6 (18%)	3	3

Predicted vs. Actual assessment outcomes taking variability into account

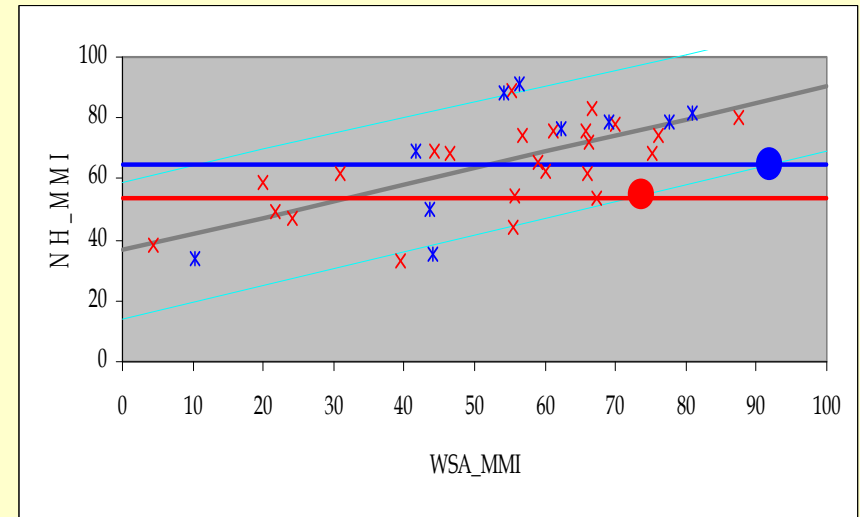
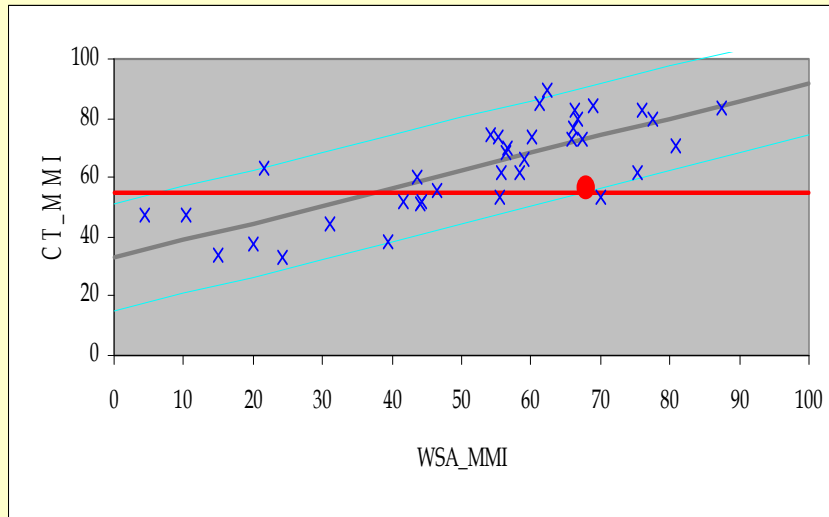


CT MMI score
necessary to achieve
NS for NH; 17 (90);
27 (75)

x Actual MMI scores
— Predicted MMI scores
— 90% prediction interval
— 75% prediction interval

CT MMI score
necessary to achieve
FS for NH; 77 (90);
68 (75)

Predicted vs. Actual assessment outcomes taking variability into account



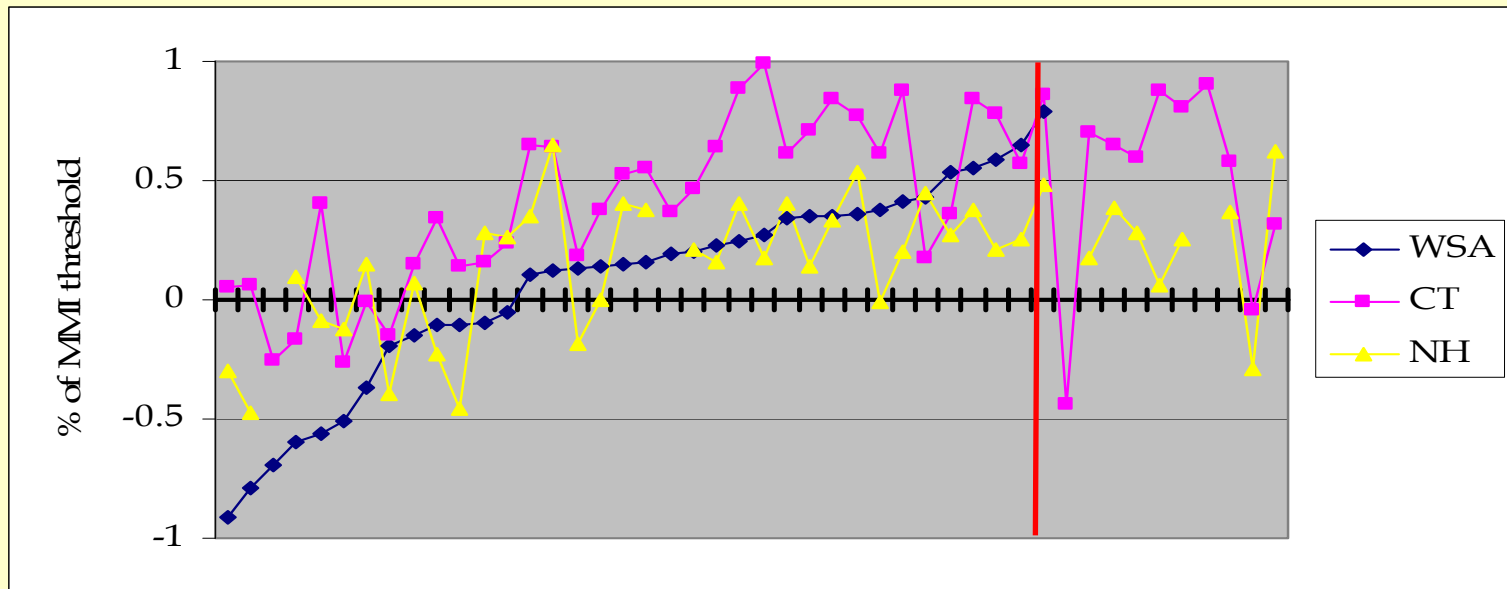
		Actual score needed for predicted assessment (90% CI)			
Actual index score	Predicted index score	FS	NS	Number possible	Number not possible
CT	NH	77	17	11	35
WSA	CT	68	(-)	7	29
WSA	NH	91/71	11/(-)	4	31

Standardizing MMI scores

Since ultimate answer is FS or NS, actual MMI relative to threshold is what matters

Standardized MMI score = (actual score - threshold) / threshold

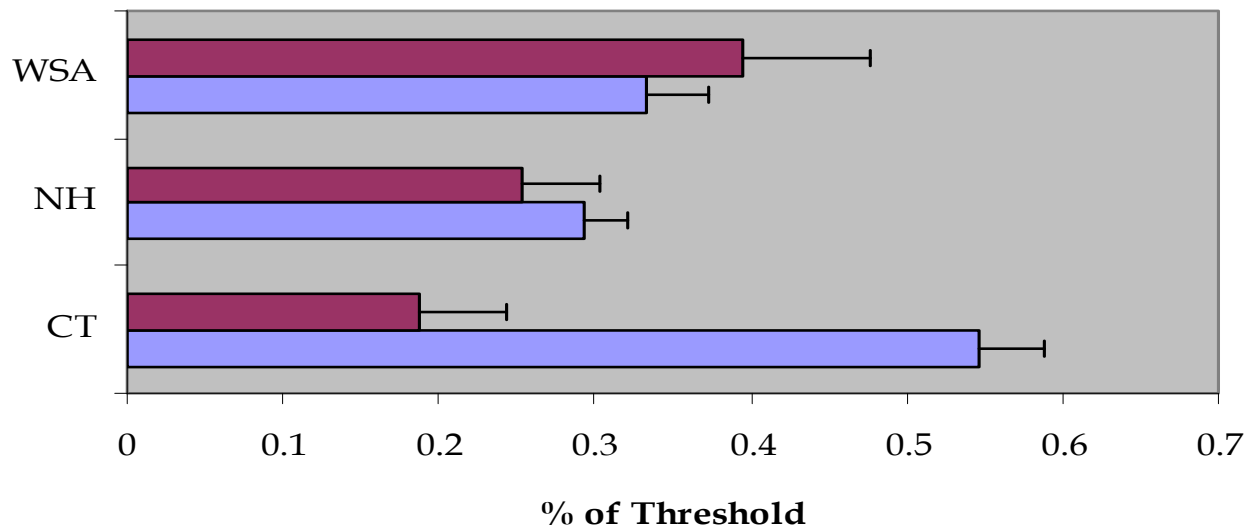
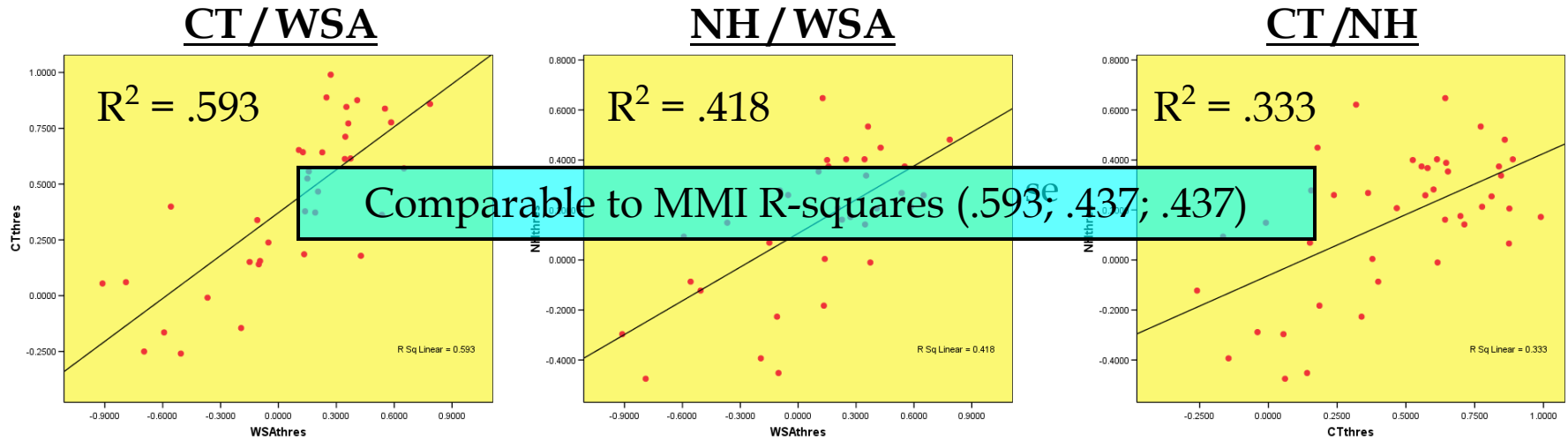
Range: -1 to +1 [(-) below threshold; (+) above threshold]



Perhaps a more informative way of looking at ALU determination than direct MMI comparisons?

Could do same “conversion” analysis with raw MMI scores

More on standardized MMI scores



WSA>NH>CT (Below) / CT<WSA<NH (Above) => WSA MMI most stringent?
Other explanations (multihabitat, etc.)

Summary

- First cut at comparing assessment outcomes for participating NE states for WSA
- On average 78% FS, 18% NS; assessment outcomes were in complete agreement 73%; NH most common dissenting opinion
- All condition indices discriminated between “Reference” and “Impacted” base on relation to non-biological indicators
- Large majority of Reference sites achieved FS; 30 – 70% of Impacted sites given NS ALU status
- ~50% variation explained by ALU status; predicted assessments agreed with actual 78%
- If take variation into account could only make assessment call using predicted MMI scores for 20% of sites; wide range of MMI scores where assessment not possible
- Standardized MMI scores may provide more informative comparison; basic comparison shows differing degrees of FS / NS achievement



Your thoughts, suggestions?

David Neils

NH Dept. Env. Services

603/271-8865

david.neils@des.nh.gov