Supplemental Material: What is the Best Wetland Assessment Tool? A Discrimination of Overlapping Assessment Approaches.

Kleindl, W. J.[1], S. P. Church[2], M. C. Rains[3], and R. Ulrich[4].

Date: February 8, 2023

Addresses:

1. Department of Land Resources and Environmental Science, Montana State University. Bozeman, MT, USA 59717 William.kleindl@montana.edu
2. Department of Earth Science, Montana State University. Bozeman, MT, USA 59717
3. School of Geosciences, University of South Florida, Tampa, FL, USA 33620
4. Los Alamos National Lab, Los Alamos, NM, USA. 87545

Keywords: Wetland, Floodplain, Riparian, Assessment, Regulatory, Factor Analysis, Qualtrics Survey, Supplemental Material

Assessment of Ecological Assessments

Last compiled on 08 February, 2023

Contents

1 Study area **1**

2 Data Visualization **1**

2.1 Overall scoring summary by tool 2

2.2 Mainly intact sites (< 2.6 LDI) 3

2.3 Mid-range sites (2.6-4.5 LDI) 9

2.4 Most disturbed sites (>4.5 LDI) 14

2.5 Summary perspective of LDI groupings 19

3 Analyses **20**

3.1 Factor Analysis 20

4 Discussion **30**

4.1 Interpretation of results 30

4.2 Limitations 31

5 Citations **32**

6 Code **34**

1 Study area

Within Southwestern Montana, 18 wetland area sites were chosen and assessed by all six tools currently used by the state of Montana for evaluating ecological health, and by the New LDI criteria. Because the sample is non-random, implications of this report cannot be extended beyond the sample to the larger area of Southwestern Montana or the state of Montana.

2 Data Visualization

Six assessment tools are used by various organizations throughout the state of Montana in an effort to characterize ecological health. Each tool was implemented at each of the 18 sites, scores for the individual ecological health indicators were standardized (scale ranging [0,1]) and then averaged to yield an overall health integrity score. Scores near one indicate nearly pristine sites, while scores near zero indicate heavily

disturbed sites. The following list details the full name of each assessment tool and the abbreviation used hereafter.

- U.S. Army Corps of Engineers, 'Hydrogeomorphic Approach' (HGM)
- Montana Department of Transportation, Montana Wetland Assessment Method (MWAM)
- Montana Natural Heritage Program, Ecological Integrity Assessment (EIA)
- United States Department of Agriculture, Natural Resources Conservation Service, Riparian Assessment Method (NRCS)
- United States Department of the Interior, Bureau of Land Management, Riparian Area Management Proper Functioning Condition Assessment for Lotic Areas (PFC)
- Montana Department of Environmental Quality (DEQ), Wetland Program Wetland Assessment Protocol (DEQ)

The 18 sites have been assessed and assigned an ecological health score ranging from [0,6] using 'New LDI' criteria. New LDI scoring functions in reverse from the normalized scoring of all six assessment tools under consideration. For all six tools (HGM, MWAM, EIA, NRCS, PFC, DEQ), lower scores indicate less-impacted, more intact sites. Higher scores indicate more disturbed, less intact sites. Higher scores on the New LDI scale indicate more pristine sites while lower scores indicate more heavily impacted sites . New LDI scores are not meant to function as 'truth' but are a useful metric for grouping sites into three general categories: mostly intact (low-scoring sites with scores less than 1), mid-range (sites scoring from 1-3), and disturbed (high-scoring sites with scores greater than 3). This grouping will be used to help assess how well various tools describe site health in each of the categories.

The following visualizations show an overview of each of the categories, giving average scores by tool for each site. A case study site is then chosen for each category, and visualizations show individual health indicator scores by tool in addition to the average scores. All visualizations and analyses were conducted in the R software environment, version 3.6.2 (R core Team, 2020).

2.1 Overall scoring summary by tool

Table **??** displays summary statistics for average site scores by each assessment tool. The BLM and NRCS tools have the highest mean average site score at 0.8627 and 0.8583, respectively. These tools also display the highest maximum average site score, giving at least one site a perfect score of one, indicating the site is in 'pristine' ecological condition. The HGM tool shows the highest range of scores (0.72) and the lowest minimum score (0.19) Both the DEQ and MTNHP tools have the most limited ranges (0.28 and 0.29, respectively) and highest minimum site scores (0.63 and 0.61, respectively).

Table 1: Summary statistics across sites by assessment tool. Average overall score, minimum overall site score, maximum overall site score, and obseved range of overall site scores for the sample.

assessment tool	mean average score	min score	max score	score range
LDI	0.58	0.23	0.87	0.64
MWAM	0.56	0.35	0.75	0.4
PFC	0.86	0.44	1	0.56
HGM	0.61	0.19	0.91	0.72
DEQ	0.77	0.63	0.91	0.28
EIA	0.78	0.61	0.9	0.29
NRCS	0.86	0.43	1	0.57

Average site scores for the entire sample by each assessment tool are displayed in Figure 1. Tool names appear on the y-axis, average site score on the x-axis, and individual points indicate the ecological condition of each site as described by a single, overall average score. These average scores are created by equally weighting all individual indices for each tool. Data points are colored by groupings of LDI scores. On the LDI scale, scores lower than 2.6 are considered 'mainly intact' (purple), scores between 2.6 and 4.5 are 'mid-range' sites (green), and those scoring above 4.5 are 'mostly disturbed' (yellow) sites.

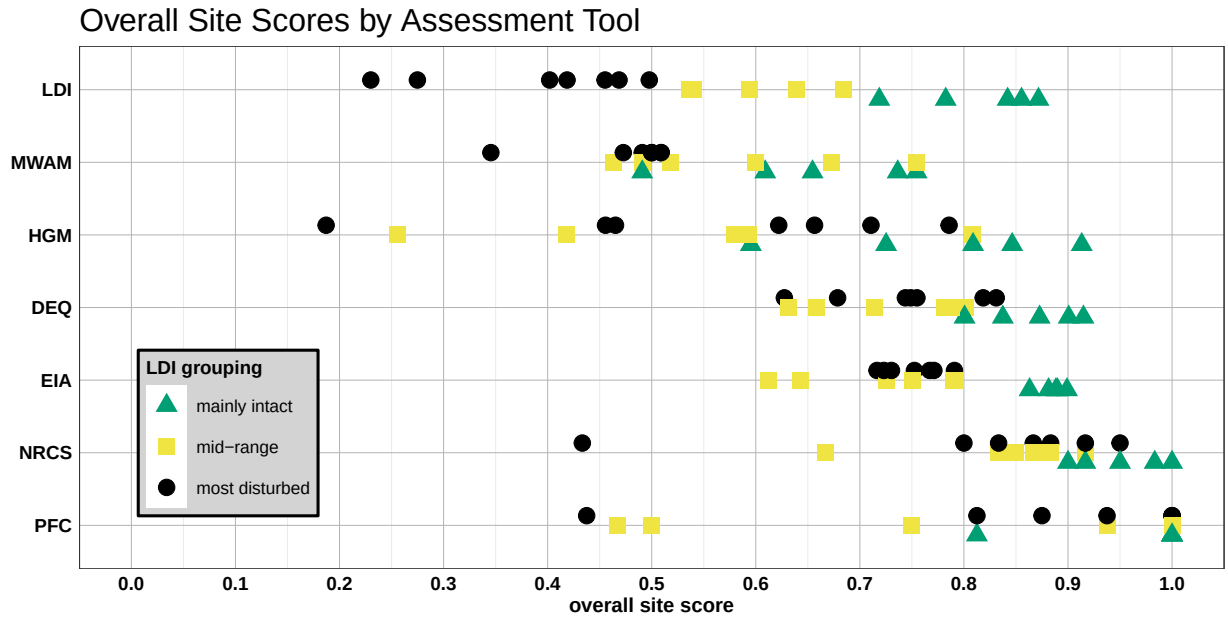


Figure 1: Average site scores for the entire sample by each assessment tool, including LDI criteria. Tool names appear on the y-axis, average site score on the x-axis, and individual points indicate the ecological condition of each site as described by a single, overall average score. These average scores are created by equally weighting all individual indices for each tool. Data points are colored by groupings of LDI scores. On the LDI scale, scores lower than 2.6 are considered 'mainly intact' (green triangles), scores between 2.6 and 4.5 are 'mid-range' sites (yellow squares), and those scoring above 4.5 are 'mostly disturbed' (black circles) sites. Data points have been vertically jittered in an effort to avoid overlapping points for tools that have assigned sites similar or identical overall scores. The PFC has assigned several sites a score of 1.0, so individual points with this score may be difficult to distinguish.

2.2 Mainly intact sites (< 2.6 LDI)

Sites 4, 6, 7, 8, and 18 comprise the mostly intact site group based upon LDI scores. Table 2 displays scores for each site by assessment tool. The PFC consistently gives one of the highest average scores for each site, scoring only one site below a perfect score of 1.0 (site 7).

Table 2: Mostly intact sites (4, 6, 7, 8, and 18) overall site scores by assessment tool. LDI scores are re-scaled to fall within the [0,1] range for comparison to other toolsThe PFC assigns a perfect score, indicating a pristine site in all regards, to all sites except site 7..

site	assessment tool	score
Site 4	LDI	0.7825
Site 4	HGM	0.9132

site	assessment tool	score
Site 4	MWAM	0.7545
Site 4	DEQ	0.9149
Site 4	EIA	0.8991
Site 4	NRCS	0.9167
Site 4	PFC	1
Site 6	LDI	0.7187
Site 6	HGM	0.7252
Site 6	MWAM	0.7364
Site 6	DEQ	0.9007
Site 6	EIA	0.8814
Site 6	NRCS	0.9833
Site 6	PFC	1
Site 7	LDI	0.8716
Site 7	HGM	0.5953
Site 7	MWAM	0.6545
Site 7	DEQ	0.8006
Site 7	EIA	0.8631
Site 7	NRCS	0.9
Site 7	PFC	0.8125
Site 8	LDI	0.8554
Site 8	HGM	0.8088
Site 8	MWAM	0.6091
Site 8	DEQ	0.8374
Site 8	EIA	0.8898
Site 8	NRCS	1
Site 8	PFC	1
Site 18	LDI	0.842
Site 18	HGM	0.8464
Site 18	MWAM	0.4909
Site 18	DEQ	0.8727
Site 18	EIA	0.8884
Site 18	NRCS	0.95
Site 18	PFC	1

Table 3 displays the mean, minimum, maximum and observed range of scores by tool for the mainly intact site grouping. The BLM and NRCS tools have the highest mean average scores at 0.96 and 0.95, and are the only two tools to assign any site a perfect score of 1. HGM shows the largest score range for this set (0.31) with one of the lower minimum scores (0.6) but a maximum score (0.9) similar to most other tools. The outlier for this grouping seems to be the MDOT tool, as it has the lowest average score (0.65), lowest minimum score (0.49), and lowest maximum score (0.75) for mainly intact sites, suggesting that for this sample it may not be capable of describing sites with fewer disturbances.

Table 3: Summary statistics across mainly intact sites by assessment tool. Average overall score, minimum overall site score, maximum overall site score, and observed range of overall site scores for the sites 4, 6, 7, 8 and 18.

assessment tool	mean average score	min score	max score	score range
LDI	1.67	1.15	2.52	1.37
MWAM	0.65	0.49	0.75	0.26
HGM	0.78	0.6	0.91	0.31
DEQ	0.87	0.8	0.91	0.11
EIA	0.88	0.86	0.9	0.04
NRCS	0.95	0.9	1	0.1
PFC	0.96	0.81	1	0.19

Figure 2 displays average scores by assessment tool for each of the five mostly intact sites. Tools are on the y-axes, score on the x-axis, and the graph is faceted by site. Re-scaled LDI scores for each site appear as black vertical bars, and site faceting appears in order of LDI site scores beginning with the lowest-scoring site in the top facet. We can see that the MDOT tool consistently gives one of the lowest average site scores across the group except for site 7, which is considered the most intact site in the sample by LDI scoring. HGM assigns its lowest score of the group to site 7 but agrees fairly well with LDI scoring for sites 6, 8 and 18. The BLM tool consistently gives one of the highest average scores for each site, scoring only one site below a perfect score of 1.0 (site 7). The MDOT and BLM tools occupy the extremes in this tier of sites and don't seem able to effectively describe disturbance for the more intact portion of the sample.

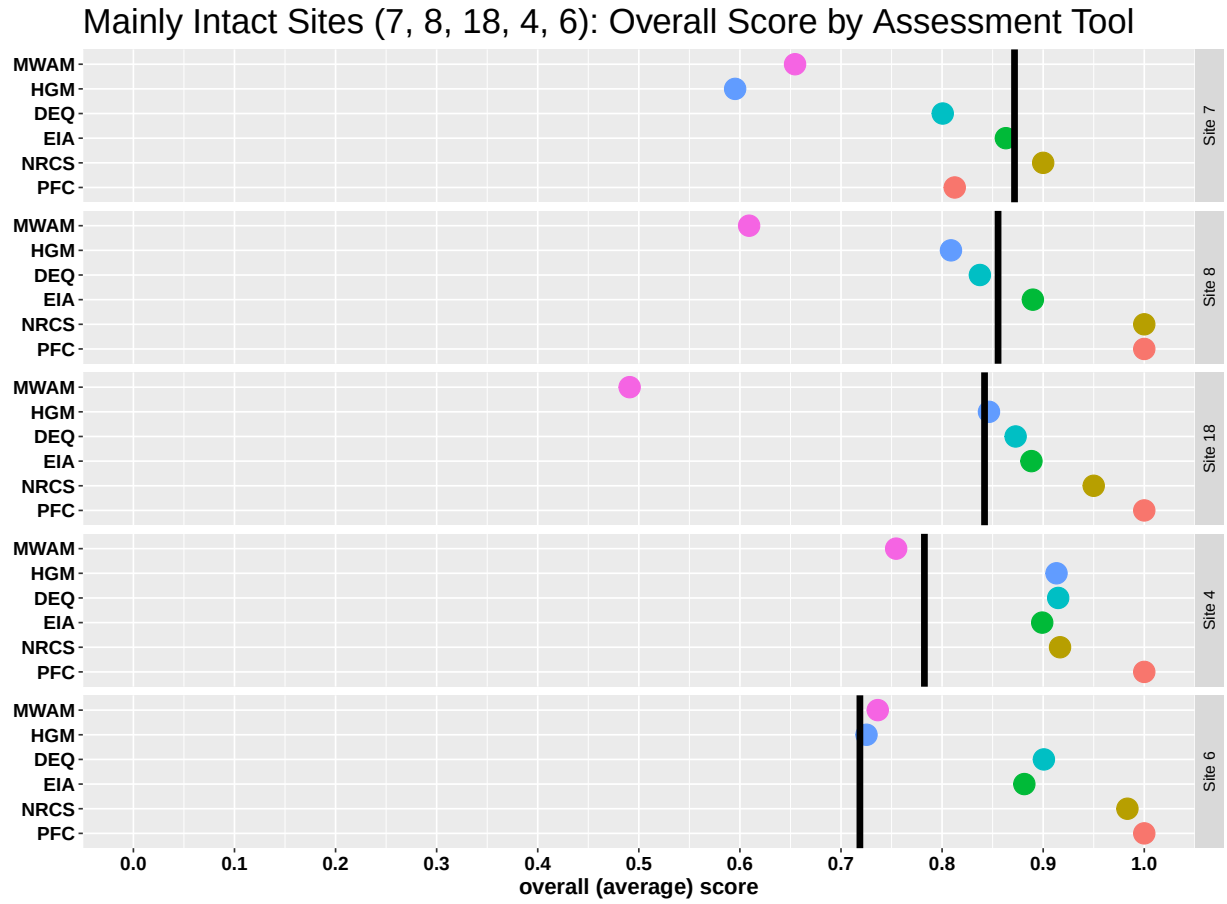


Figure 2: Comparison of average site scores (x-axis) for mostly intact sites by assessment tool (y-axes). Colored by tool and faceted by site number. Rescaled LDI scores for each site appear as black vertical bars, and site faceting appears in order of LDI site scores beginning with the highest-scoring LDI site in the top facet.

2.2.1 Case Study: Site 7, the most intact of the mostly intact sites by LDI criteria

Site 7 is located 25 miles north of Ennis, MT, along the Madison River immediately south of Cherry Creek in Bear Trap Canyon. This site scores a 1.15 by LDI criteria and should be considered one of the most intact sites in the sample. Figure 3 looks at the health indicator categories in comparison to the overall score for each of the assessment tools at this site.

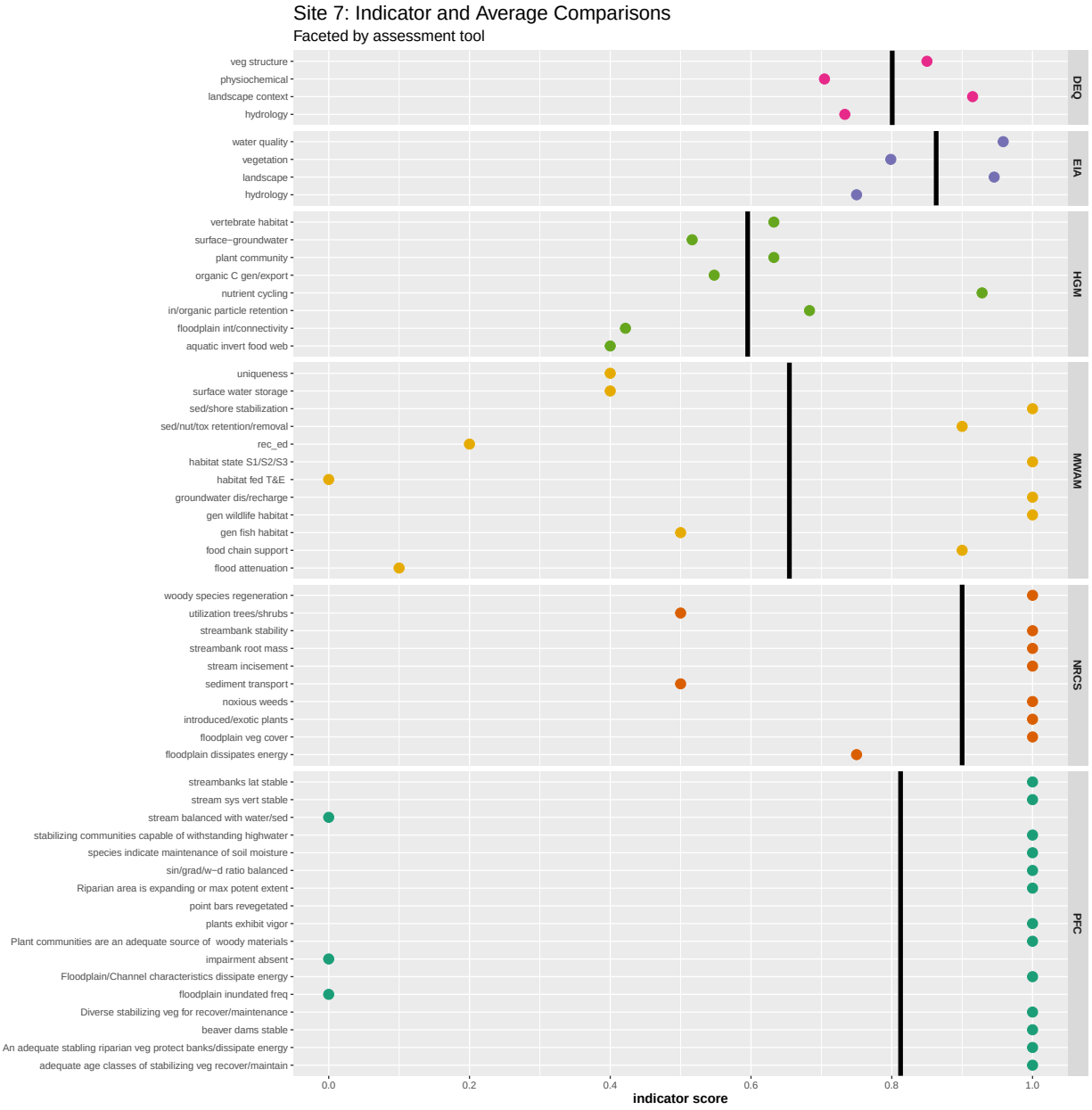


Figure 3: Comparison of site 7 habitat indicator scores (individual dots) and averages (bold vertical line) for each assessment tool. Individual indicator scores are on the y-axis, score on the x-axis, faceted and colored by tool. The NCRS tool scores this site the highest overall, while the HGM tool scores this site the lowest.

2.3 Mid-range sites (2.6-4.5 LDI)

Sites 3, 9, 10, 12, 13, and 17 comprise the mid-range site group based upon LDI scores. Table 4 displays the average scores by tool for each of the seven sites scoring between a 2.6 and 4.5 based on LDI criteria.

Table 4: Mid-range sites (3, 9, 10, 12, 13, and 17) overall site scores per assessment tool. LDI scores are re-scaled to fall within the [0,1] range for comparison to other tools.

site	assessment tool	score
Site 3	LDI	0.5403
Site 3	HGM	0.808
Site 3	MWAM	0.6727
Site 3	DEQ	0.8013
Site 3	EIA	0.7903
Site 3	NRCS	0.9167
Site 3	PFC	0.9375
Site 9	LDI	0.6389
Site 9	HGM	0.2556
Site 9	MWAM	0.5182
Site 9	DEQ	0.6312
Site 9	EIA	0.6432
Site 9	NRCS	0.8333
Site 9	PFC	0.4667
Site 10	LDI	0.5363
Site 10	HGM	0.5927
Site 10	MWAM	0.4909
Site 10	DEQ	0.7812
Site 10	EIA	0.7251
Site 10	NRCS	0.8667
Site 10	PFC	1
Site 12	LDI	0.594
Site 12	HGM	0.4181
Site 12	MWAM	0.6
Site 12	DEQ	0.6582
Site 12	EIA	0.6119
Site 12	NRCS	0.6667
Site 12	PFC	0.5
Site 13	LDI	0.6846
Site 13	HGM	0.5793
Site 13	MWAM	0.4636
Site 13	DEQ	0.7879
Site 13	EIA	0.7915
Site 13	NRCS	0.85
Site 13	PFC	0.75
Site 17	LDI	0.6389
Site 17	HGM	0.592
Site 17	MWAM	0.7545
Site 17	DEQ	0.7145
Site 17	EIA	0.7507
Site 17	NRCS	0.8833
Site 17	PFC	1

Table 5 displays the mean, minimum, maximum and observed range of scores by tool for the mid-range site grouping. The PFC has the highest maximum score of one, and is the only tool to describe any mid-range site as 'pristine'. The PFC and HGM Approach have the largest range of scores (0.53 and 0.55, respectively), with the HGM range shifted to the lower end of scoring, extending from 0.26 (the lowest minimum score) to 0.81. Although the NRCS does not describe any of the mid-range sites as 'pristine', it exhibits consistently higher scores overall with the highest mean average site score (0.84) and the highest minimum (0.67). The DEQ and EIA have the most limited observed scoring ranges (0.17 and 0.18, respectively), and seem to behave similarly for this tier of sites. The minimum, maximum, and mean average scores are quite similar and do seem to consistently assign middle-of-the-range scores for this tier of sites.

Table 5: Summary statistics across mid-range sites by assessment tool. Average overall score, minimum overall site score, maximum overall site score, and observed range of overall site scores for the sites 3, 9, 10, 12, 13, and 17.

assessment tool	mean average score	min score	max score	score range
LDI	3.54	2.83	4.16	1.33
MWAM	0.58	0.46	0.75	0.29
HGM	0.54	0.26	0.81	0.55
DEQ	0.73	0.63	0.8	0.17
EIA	0.72	0.61	0.79	0.18
NRCS	0.84	0.67	0.92	0.25
PFC	0.78	0.47	1	0.53

Figure 4 displays average scores by assessment tool for each of the six mid-range sites. Tools are on the y-axes, score on the x-axis, and the graph is faceted by site. Re-scaled LDI scores for each site appear as black vertical bars, and site faceting appears in order of LDI site scores beginning with the lowest-scoring site in the top facet. The PFC gives the highest scores for three of the sites (10, 3, 17) including describing two of these sites as 'pristine' with a perfect score of 1. Scores for the other sites are some of the lowest across all tools and the general scoring of this tier of sites does not line up well with LDI scoring. The NRCS is consistent in assigning one of the highest scores across all mid-range sites and does not seem able to capture disturbance for this tier of sites. The DEQ and EIA score all of the mid-range sites very similarly and agree fairly well with LDI scoring. All of the tools seem to diverge from LDI scoring for site 3. For sites 12, 9 and 17, the HGM Approach disagrees with not only LDI scoring but the scoring of all of the other five tools, assigning the lowest score for all three sites.

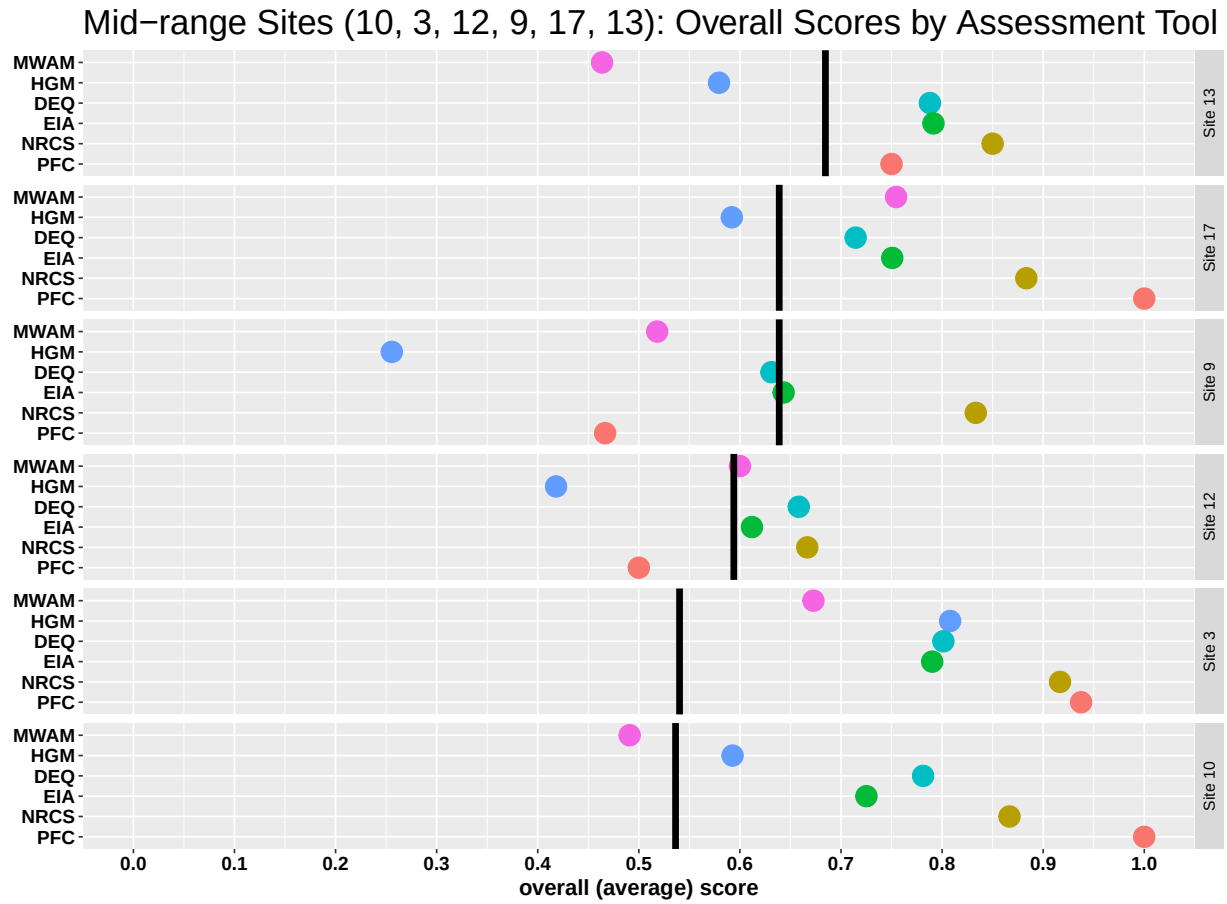


Figure 4: Comparison of average site scores (x-axis) for mid-range sites by assessment tool (y-axes). Colored by tool and faceted by site number. Rescaled LDI scores for each site appear as black vertical bars, and site faceting appears in order of LDI site scores beginning with the lowest-scoring LDI site in the top facet.

2.3.1 Case Study: Site 3, a mid-range site (4.12 LDI)

Site 3 (LDI score of 4.12) is located along Middle Cottonwood Creek near Riverside Country Club in northwest Bozeman off Springhill Road. Figure 5 looks at the health indicator categories in comparison to the overall score for each of the assessment tools at this site.

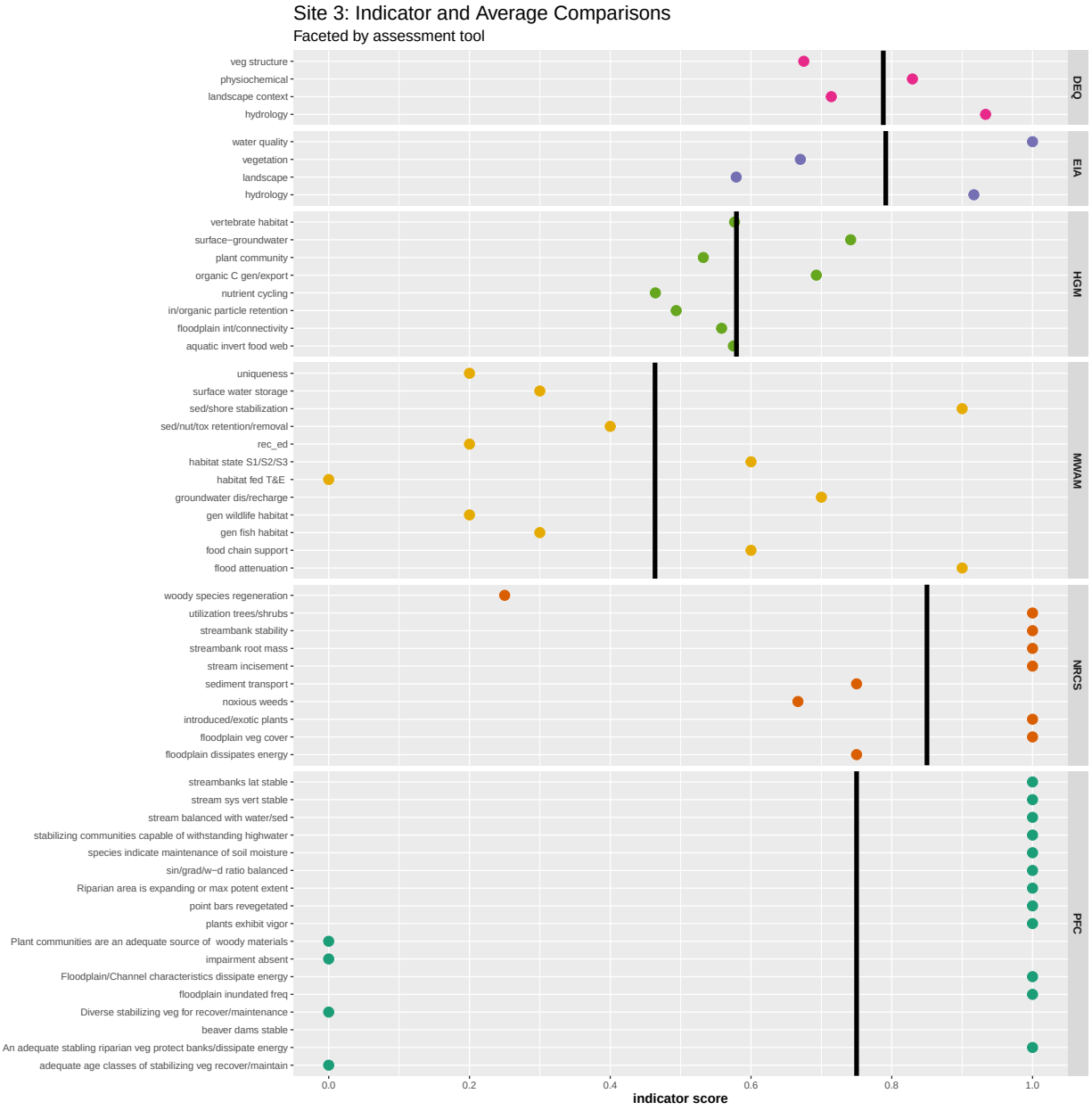


Figure 5: Comparison of site 3 habitat indicator scores (individual dots) and averages (bold vertical line) for each assessment tool. Individual indicator scores are on the y-axis, score on the x-axis, faceted and colored by tool. The NCRS tool scores this site the highest overall, while the MDOT tool scores this site the lowest.

2.4 Most disturbed sites (>4.5 LDI)

Sites 1, 2, 5, 11, 14, 15 and 16 comprise the most disturbed group based upon 'New LDI' scores. Table 6 displays summaries of the average scores by tool for each of the seven sites scoring above a 4.5 based on 'New LDI' criteria.

Table 6: Mostly disturbed sites (1, 2, 5, 11, 14, 15 and 16) overall site scores by assessment tool. LDI scores are re-scaled to fall within the [0,1] range for comparison to other tools. The PFC assigns several perfect scores to these sites, indicating a lack of ability to detect sites with compromised ecological integrity.

site	assessment tool	score
Site 1	LDI	0.4551
Site 1	HGM	0.7857
Site 1	MWAM	0.4909
Site 1	DEQ	0.8311
Site 1	EIA	0.7711
Site 1	NRCS	0.8
Site 1	PFC	1
Site 2	LDI	0.402
Site 2	HGM	0.7107
Site 2	MWAM	0.5091
Site 2	DEQ	0.8185
Site 2	EIA	0.7909
Site 2	NRCS	0.8667
Site 2	PFC	1
Site 5	LDI	0.2301
Site 5	HGM	0.1871
Site 5	MWAM	0.3455
Site 5	DEQ	0.6275
Site 5	EIA	0.7165
Site 5	NRCS	0.4333
Site 5	PFC	0.4375
Site 11	LDI	0.4686
Site 11	HGM	0.4652
Site 11	MWAM	0.5091
Site 11	DEQ	0.7549
Site 11	EIA	0.7524
Site 11	NRCS	0.95
Site 11	PFC	1
Site 14	LDI	0.4187
Site 14	HGM	0.6221
Site 14	MWAM	0.5
Site 14	DEQ	0.7438
Site 14	EIA	0.7304
Site 14	NRCS	0.9167
Site 14	PFC	0.875
Site 15	LDI	0.4977
Site 15	HGM	0.6565
Site 15	MWAM	0.4727
Site 15	DEQ	0.7489
Site 15	EIA	0.7671
Site 15	NRCS	0.8833

site	assessment tool	score
Site 15	PFC	0.9375
Site 16	LDI	0.2748
Site 16	HGM	0.4556
Site 16	MWAM	0.5
Site 16	DEQ	0.6786
Site 16	EIA	0.7231
Site 16	NRCS	0.8333
Site 16	PFC	0.8125

Table 7 displays the mean, minimum, maximum and observed range of scores by tool for the most disturbed site grouping. The BLM tool has the highest maximum score of one - the only tool to describe any mostly disturbed site as 'pristine' - and also has the highest mean overall site score. This is closely followed by the NRCS tool with the second-highest maximum (0.95) and mean average site score (0.81). The MDOT tool exhibits the lowest mean average score (0.48), the lowest minimum score (0.35) and lowest maximum score (0.51). THe MTNHP tool has a very small range of observed scores, showing a difference of only 0.07 between the minimum and maximum score for this grouping. The HGM tool has the lowest minimum score (0.19) and the highest range of observed scores (0.6) for this tier of sites.

Table 7: Summary statistics across the most disturbed sites by assessment tool. Average overall score, minimum overall site score, maximum overall site score, and observed range of overall site scores for the sites 1, 2, 5, 11, 14, 15 and 16.

assessment tool	mean average score	min score	max score	score range
LDI	5.45	4.51	6.91	2.4
MWAM	0.48	0.35	0.51	0.16
HGM	0.55	0.19	0.79	0.6
DEQ	0.74	0.63	0.83	0.2
EIA	0.75	0.72	0.79	0.07
NRCS	0.81	0.43	0.95	0.52
PFC	0.87	0.44	1	0.56

Figure 6 display average scores by assessment tool for each of the seven most disturbed sites. Tools are on the y-axes, score on the x-axis, and the graph is faceted by site. Rescaled LDI scores for each site appear as black vertical bars, and site faceting appears in order of LDI site scores beginning with the lowest-scoring site in the top facet. The BLM tool gives three of the seven sites a perfect score (sites 2, 1, and 11), and along with NRCS tool assigns high scores of 0.8 or above across all sites with the exception of site 5. This indicates that for this sample, these two tools are generally unable to capture disturbance for this tier of sites, or to differentiate the more disturbed from the least disturbed sites. Although the MDOT consistently gives one of the lower scores of all the tools and is relatively well-aligned with LDI scoring. For six of the seven sites it shows little to no difference in scoring, which is somewhat mirrored by LDI scoring, exhibiting a limited range for sites 2, 14, 1, 11, and 15. The MTNHP tool scores all of the most disturbed sites between 0.7 and 0.8 indicating that not only is it unable to recognize more disturbed sites in this sample, it is also unable to capture nuance of disturbance within this tier of sites. The HGM tool does the best job of capturing disturbance for the most disturbed site in the sample (site 5), and is reasonably well-aligned with LDI scoring for sites 16, 14, 11 and 15. However, this tool's mean average score for the most disturbed sites (0.55) is quite similar to the mean average score for the mid-range sites (0.54), suggesting that these two tiers of sites have no clear distinction within the HGM scoring system.

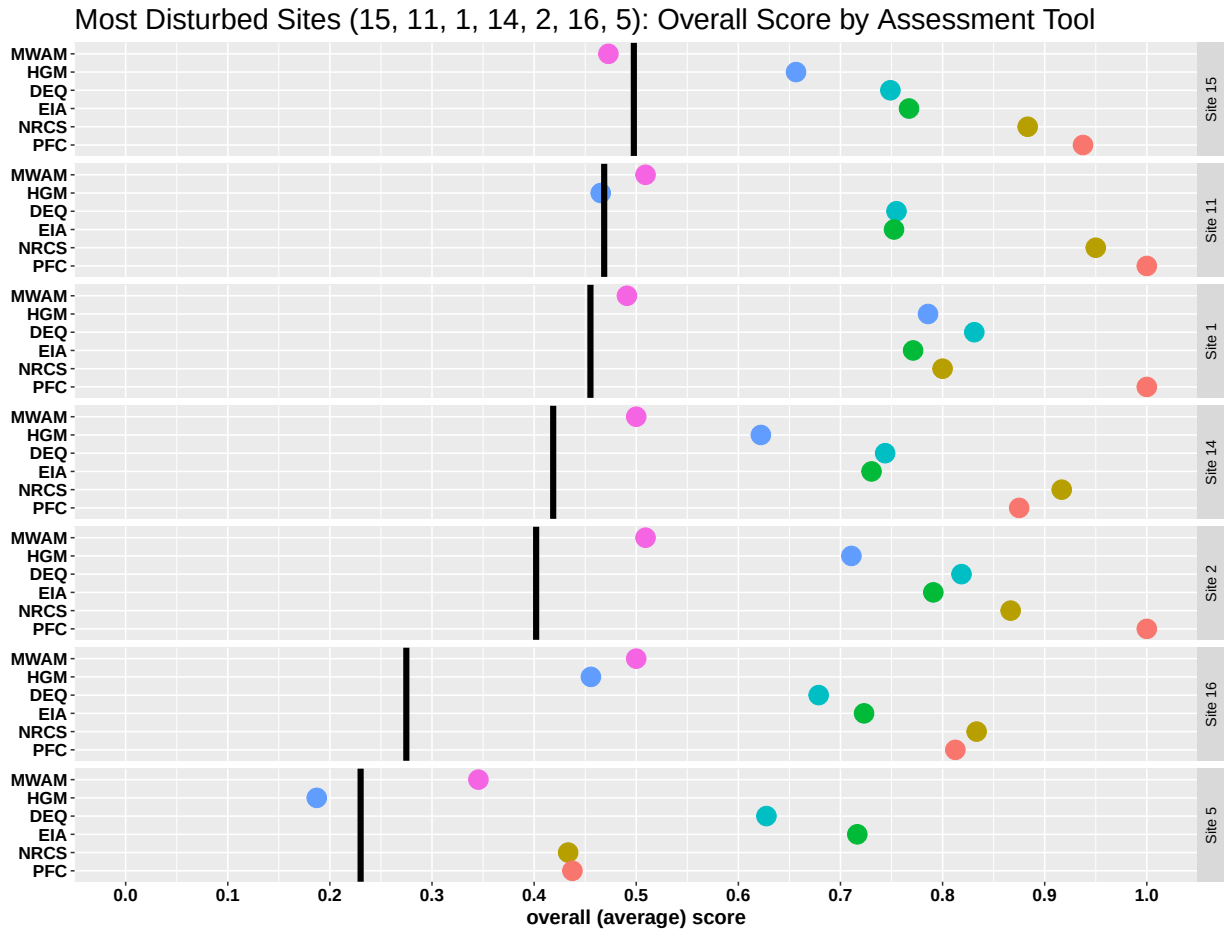


Figure 6: Comparison of average site scores (x-axis) for most disturbed sites by assessment tool (y-axes). Colored by tool and faceted by site number. Rescaled LDI scores for each site appear as black vertical bars, and site faceting appears in order of LDI site scores beginning with the lowest-scoring LDI site in the top facet.

2.4.1 Case Study: Site 5, the most disturbed site (6.9 LDI)

Site 5 has a score of 6.9 according to 'New LDI' criteria, and is the highest-scoring site in the sample, indicating that it should be considered one of the more disturbed sites by all assessment tools. The site 5 wetland area is located along the Jefferson River, just south of the confluence with the Big Hole River, in Twin Bridges of Madison County, Montana. To the east lies Old Baldy Mountain in Beaverhead-Deerlodge National Forest and to the west is McCartney Mountain. Figure 7 looks at the health indicator categories in comparison to the overall score for each of the assessment tools at this site.

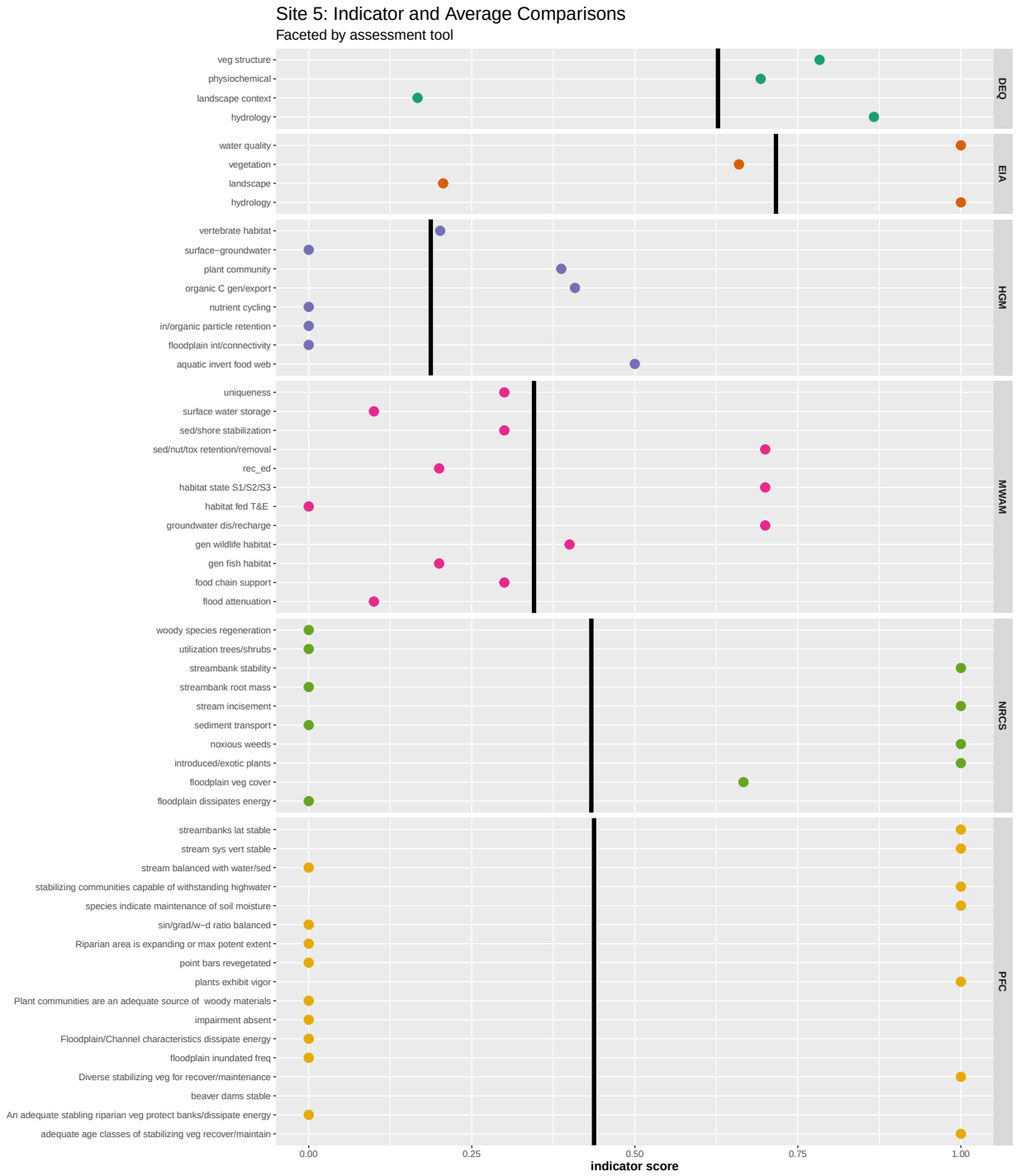


Figure 7: Comparison of site 5 habitat indicator scores (individual dots) and averages (bold vertical line) for each assessment tool. Individual indicator scores are on the y-axis, score on the x-axis, faceted and colored by tool. The HGM tool scores this site the lowest overall, while the MTNHP tool scores this site the highest.

2.5 Summary perspective of LDI groupings

Figure 8 gives a summary perspective of the sites as grouped by LDI scores. The figure is faceted by LDI groupings, labeled on the right-hand side (mainly intact, mid-range, most disturbed), and organized on the y-axis by tool. Individual data points represent the average score for a single site as determined by each tool. Points are colored by their corresponding LDI score with yellow-orange hues indicating scores closer to zero and purple-black hues indicating scores closer to one (color legend appears inset in the bottom, left-hand corner of the graph). Points are labeled by site number with indicator arrows used for readability when data points overlap.

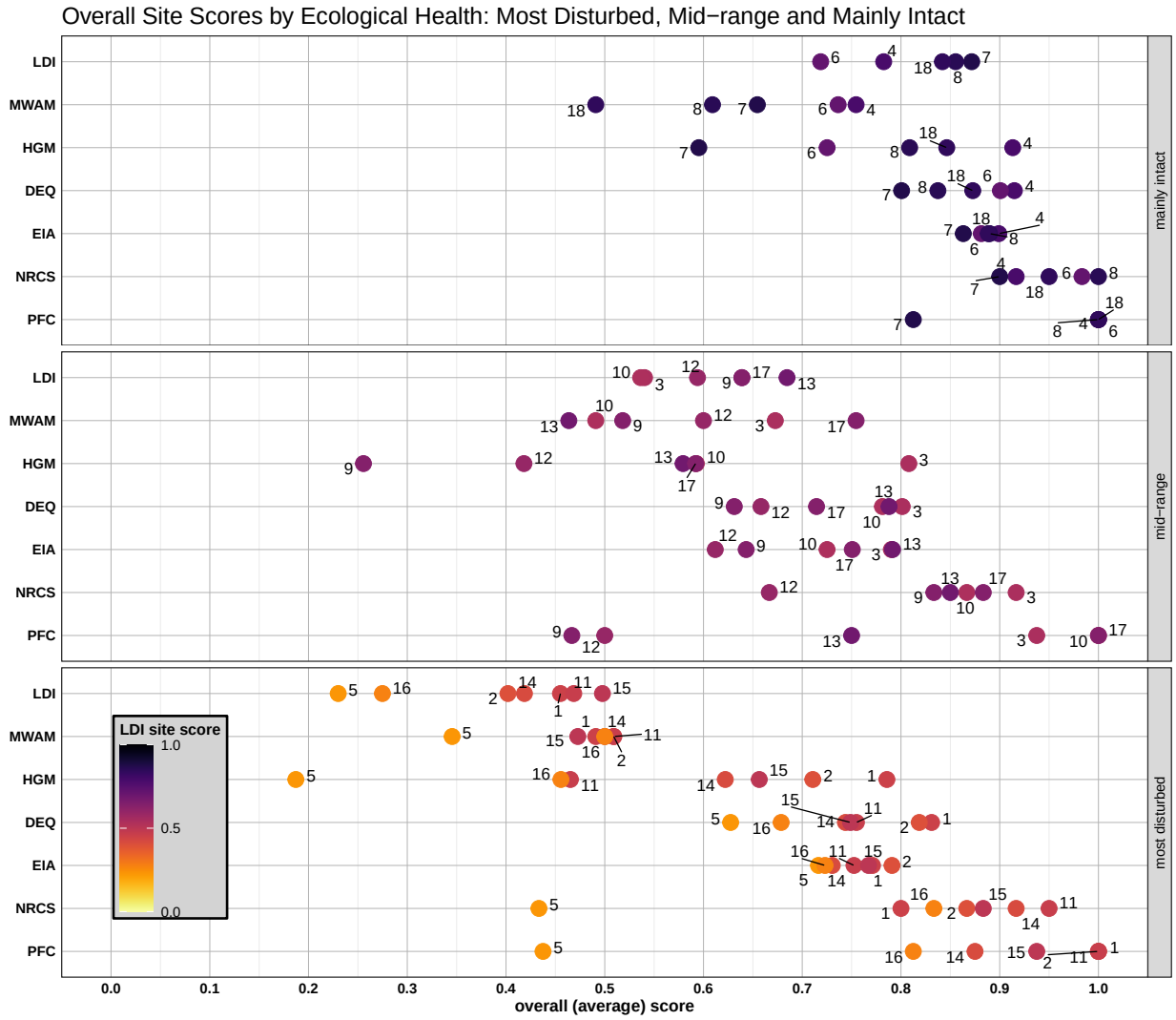


Figure 8: Summary perspective of the sites as grouped by LDI scores. Faceted by LDI groupings, labeled on the right-hand side (mainly intact, mid-range, most disturbed), and organized on the y-axis by tool. Individual data points represent the average score for a single site as determined by each tool. Points are colored by their corresponding LDI score and labeled by site number. Scores approaching zero are more disturbed (yellow-orange hues) while scores approaching one are less disturbed (purple-black hues).

3 Analyses

3.1 Factor Analysis

The goal in implementing each of the assessment tools is to describe or measure the concept of 'ecological health'. Ecological health can be thought of as a latent variable because it cannot be directly measured. Instead, we arrive at a statistic to represent this idea by averaging individual health factors. Individual health factors represent the various aspects of ecological health that can be measured, and vary in number and description for each assessment tool. We assume that each of the health factors contribute equally to describing the single underlying concept of ecological health. Factor analysis can help to assess if this assumption is reasonable:

"In parsimoniously describing data, factor analysts explicitly recognize that any relationship is limited to a particular area of applicability. Areas qualitatively different, that is, areas where relatively little generalization can be made from one area to another, are referred to as *separate factors*. Each factor represents an area of generalization that is qualitatively distinct from that represented by any other factor. Within an area where data can be summarized, factor analysts first represent that area by a factor and then seek to make the degree of generalization between each variable and the factor explicit. A measure of the degree of generalizability found between each variable and each factor is calculated and referred to as a *factor loading*. Factor loadings reflect quantitative relationships. The farther the factor loading is from zero, the more one can generalize from that factor to the variable" (Gorsuch, 1983).

3.1.1 Exploratory factor analysis

Exploratory factor analysis (EFA) is "an attempt to approximate a correlation or covariance matrix with one of less rank" (Revelle, 2019). EFA is "focused on modeling the entire off-diagonal covariance matrix, especially the covariances between the variables, so standardizing the variables does not impact the results" (Greenwood, 2020). This is important because all of the tools have their own scoring system and scale, so in order to compare tools all of the scales were standardized to exist between [0,1]. We are mainly interested in assessing the construct validity of a tool's one-number summary of a given site. That is, do the multi-metric indices actually measure the theorized concept of 'ecological health'?

Why EFA and not confirmatory factor analysis (CFA)? EFA is typically used to discover the underlying structure and CFA is used to test a theory about the underlying structure in a dataset. "Computationally, CFA is much more 'expensive' than EFA and some data sets where EFA is possible will not provide estimable models in CFA" (Greenwood, 2020). We have a limited sample size, which reflects the state of gathering data in wetland ecology. It is time-consuming and expensive to assess wetlands with a single tool, let alone multiple tools, thus designing a procedure for dealing with a limited sample size is important in practice. For this analysis, some tools already necessitate a variable reduction procedure to use an EFA approach so it seems likely that a CFA approach will not be practical for this situation. If the data do not seriously violate the assumptions of multivariate normality, then a maximum likelihood approach to EFA can be used, which offers some advantages that encourage the use of EFA in this situation.

"We suggest that factor analysis is preferable to principle components analysis. Components analysis is only a data reduction method. It is computed without regard to any underlying structure caused by latent variables; components are calculated using all of the variance of the manifest variables, and all of that variance appears in the solution. However, researchers rarely collect and analyse data without an *a priori* idea about how the variables are related. The aim of factor analysis is to reveal any latent variables that cause the manifest variables to covary. During factor extraction the shared variance of a variable is partitioned from its unique variance and error variance to reveal the underlying factor structure; only shared variance appears in the solution. Principal components analysis does not discriminate between shared and unique variance. When the factors are uncorrelated and communalities are moderate it can produce inflated values of variance accounted for by the components. Since factor analysis only analyzes shared variance, factor analysis should yield the same solution (all other things being equal) while also avoiding the inflation of

estimates of variance accounted for" (Costello and Osborne, 2005). Due to the small sample size, we intend to use Ordinary Least Squares (OLS) to find the minimum residual (minres) solution because "this produces solutions very similar to maximum likelihood even for badly behaved matrices" (Revelle, 2019).

3.1.2 Assessing multivariate normal assumption

"Maximum likelihood is regarded, by statisticians at least, as perhaps the most respectable method of estimating the parameters in the factor analysis. The essence of this approach is to assume that the data being analysed have a multivariate normal distribution" (Everitt & Hothorn, 2011).

3.1.3 The k-factor analysis model

"Problems... arise out of the factor-to-variable relationship. The only assumption made in using the linear model is that the variables are linear functions of the factors. The variables are not assumed to be linearly related. Variables may be related in any manner whatsoever so long as that interrelationship is a function of each variable being linearly related to the factors." (Gorsuch, 1983)

From Chapter 5 of *An Introduction to Applied Multivariate Analysis with R* (Everitt & Hothorn, 2011):

Generally, $\mathbf{x} = \mathbf{\Lambda f} + \mathbf{u}$, where:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_q \end{bmatrix}, \mathbf{\Lambda} = \begin{bmatrix} \lambda_{11} & \lambda_{12} & \dots & \lambda_{1k} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{q1} & \lambda_{q2} & \dots & \lambda_{qk} \end{bmatrix}, \mathbf{f} = \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_k \end{bmatrix}, \mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_q \end{bmatrix}$$

and, $i = 1, ..., q$ and $j = 1, ..., k$.

Our observed/manifest variables, $\mathbf{x}$, of length q, are assumed to be linked to k unobserved latent variables, $\mathbf{f}$, where $k < q$. The factor loadings, $\mathbf{\Lambda}$, show how each observed variable, x_i depend on the common factors and function like regression coefficients. The random disturbance terms, $\mathbf{u}$, are specific to each x_i, and as such are referred to as *specific variates*.

Assumptions:

- Factors are independent of one another, that is $Cov(f_j, f_{j'}) = 0$.
- Factors are unobserved, so the scale and location can be fixed arbitrarily. They are standardized to have a mean of 0 and a variance of 1.
- The specific variates, u_i, are made of two parts - the model's inability to explain the response and the fact that the observed variable is only an approximate measure of the underlying latent variable. "In practise no attempt is made to disentangle these two parts."
  - The u_i, are independent of each other, $Cov(u_i, u_{i'}) = 0$.
  - The u_i are independent of the factors, $Cov(f_j, u_i) = 0$.
  - $Var(u_i) = \psi_i$.
- Given these assumptions, show $Var(X_i) = \sigma_i^2 = \sum_{j=1}^{k} \lambda_{ij}^2 + \psi_i$:
  - Recall, $X_i = \lambda_{ij} * f_j + u_i$.
  - Then,

$$\begin{aligned} Var(X_i) &= Var(\lambda_{i1} * f_1 + \lambda_{i2} * f_2 + ... + \lambda_{ik} f_k + u_i) \\ &= Var(\lambda_{i1} * f_1) + Var(\lambda_{i2 * f_2}) + ... + Var(\lambda_{ik} * f_k) + Var(u_i) \end{aligned} \tag{1}$$

    because of properties of variance.

- $Var(X_i) = \lambda_{i1}^2 + \lambda_{i2}^2 + ... + \lambda_{ik}^2 + \psi_i$ because $Var(\lambda_{i1} * f_1) = \lambda_{i1}^2 * Var(f_1) = \lambda_{i1}^2$.
- $Var(X_i) = \sigma_i^2 = \sum_{j=1}^k \lambda_{ij}^2 + \psi_i$

- Thus, $h_i^2 = \sum_{j=1}^k \lambda_{ij}^2$ is referred to as the *communality* of the variable and represents the variance shared with the other variables via the common factors. The specific variance, ψ_i, relates to the variability in x_i not shared with other variables (not explained by the model).
- Relationships between the observed variables are a result of the relationships of these variables to the latent variables (common factors). Thus the covariance between any two observed variables is through the common factors.
  - $var(X_i, X_{i'}) = \sigma_{ii'} = \sum_{j=1}^k \lambda_{ij} * \lambda_{i'j}$
  - Implying the population covariance matrix: $Var(\mathbf{X}) = \mathbf{\Sigma} = \mathbf{\Lambda\Lambda^T} + \mathbf{\Psi}$ where $\mathbf{\Psi} = \mathbf{diag}(\psi_\mathbf{i}) = \begin{bmatrix} \psi_1 & 0 & \dots & 0 \\ 0 & \psi_2 & \vdots & \vdots \\ \vdots & \dots & \ddots & \vdots \\ 0 & \dots & 0 & \psi_\alpha \end{bmatrix}$.

3.1.4 Estimating parameters in EFA: A simple approach

An estimate for the population covariance matrix, $\mathbf{\Sigma}$, is the sample covariance matrix:

$$\mathbf{S} \approx \hat{\mathbf{\Lambda}}\hat{\mathbf{\Lambda}}^\mathbf{T} + \hat{\mathbf{\Psi}} \implies \mathbf{S} - (\hat{\mathbf{\Lambda}}\hat{\mathbf{\Lambda}}^\mathbf{T} + \hat{\mathbf{\Psi}})$$

Recall, $\hat{h_i^2} = diag(\hat{\Lambda}\hat{\Lambda}^T)$ and $\hat{\psi_i} = Var(X_i) - h_i^2 = 1 - h_i^2$ where the h_i^2 are the *communality* of the variable and represent the variance shared with the other variables via the common factors.

Thus, "the specific variances are trivially estimated to make the model perfectly reconstruct the variances of the observed variables."

$$\hat{\psi_i} = S_i^2 - \sum_{j=1}^k \hat{\lambda}_{ij}^2 = S_i^2 - \hat{h_i^2}$$

3.1.5 Rotations and transformations

If the solution is complex or the loadings do not match expectations, a rotation could be considered. Orthogonal rotations restrict the rotated factors to being uncorrelated. If it isn't reasonable to assume and force the latent traits to be uncorrelated, a transformation could be considered. A transformation is often called an oblique rotation but because the factors are allowed to be correlated, it is actually a transformation (Greenwood, 2020).

There is no unique solution in EFA - that is, there are many factor loadings that could satisfy the model:

Consider introducing an orthogonal matrix MM^T into the basic EFA model:

$$\mathbf{X} = \mathbf{\Lambda M M^T} + \mathbf{u} = (\mathbf{\Lambda M})(\mathbf{M^T f}) + \mathbf{u} = \mathbf{\Lambda M f^*} + \mathbf{u}$$

with new loadings $\Lambda^* = \Lambda M$ and new factors defined $f* = M^T f$.

Thus, the new model for the population covariance matrix is the same as the original model:

$$\mathbf{\Sigma} = \mathbf{\Lambda M M^T \Lambda} + \mathbf{\Psi} = \mathbf{\Lambda\Lambda^T}$$

"Rotation methods search for specific alternative loading patterns using a choice of M for the k factors that explain the same proportion of the total information but are hopefully easier to interpret. Sometimes also allowing them to be dependent which are called *transformations*, not *rotations*, but still define a new M" (Greenwood, 2020).

In the literature, oblique rotation and transformation are often used interchangeably, just as rotation is often substituted for orthogonal rotation. "Rotation cannot improve the basic aspects of the analysis, such as the amount of variance extracted from the items. Varimax, quartimax, and equamax are commonly available orthogonal methods of rotation; direct oblimin, quartimin, and promax are oblique. Orthogonal rotations produce factors that are uncorrelated; oblique methods allow the factors to correlate... using orthogonal rotation results in a loss of valuable information if the factors are correlated, and oblique rotation should theoretically render a more accurate, and perhaps more reproducible, solution. If the factors are truly uncorrelated, orthogonal and oblique rotation produce nearly identical results" (Costello and Obsborne, 2005).

In the *psych* package the default rotation uses an oblimin transformation, which aims to "find simple structure in the factor pattern matrix while allowing some degree of correlation between the factors" (Greenwood, 2020). The option 'promax' is the most commonly used oblique transformation. By taking a varimax rotation and transforming it by raising it to a power, a slightly simpler but oblique solution is reached by pushing loadings that were small to 0. This option looks for the best structure with the lowest correlation between factors (Greenwood, 2020). In our case, the overall average score for each tool is the sole measure of 'ecological health'. By averaging the individual health indicator groups we are assuming that a single factor exists, and this analysis seeks to assess how well that assumption is met. Therefore, as all analyses force a single factor, transformations are not meaningful in this context.

3.1.6 Application

This analysis is intended to serve as a general guideline for assessing the ability of various tools to describe a set of sample sites. Because each organization developed tools to meet their needs, there may be no single 'best' tool, and the usefulness of each tool in describing a sample varies across time and space. However, this analytical approach helps evaluators to assess which tool most efficiently captures the nuance of disturbance within the sample under consideration.

Consider that the goal of each assessment tool is to provide some measure of 'ecological condition' or the degree the system diverges from an ideal state. Here we consider the ecological condition as a latent variable because it cannot be directly measured. Instead, each tool uses metrics to measure various aspects of ecological condition (i.e., characteristic plant community, sediment/shoreline stabilization, etc.) and eventually combines these scores to yield a single index score that describes ecological condition. In combining the individual metric scores to create an overall average score, each tool assumes that all metrics are describing the single underlying concept of ecological condition. Using an exploratory factor analysis approach (Everitt and Hothorn 2011), we seek to assess if this assumption is reasonable for the sample under consideration.

All data visualizations and analyses were conducted in the R software environment version 4.0.1 (R Core Team 2020), and the accompanying R scripts can be found at the end of this document. Using the *psych* package (Revelle, 2019), factor loadings were estimated using the least squares minimum residuals (minres) algorithm. Because we are interested in evaluating the assumed underlying structure of each tool, the number of factors is forced to one, rendering transformations meaningless in this context. Data visualization, in the form of tile plots (Wickham et al., 2019), graphic representations of correlation coefficient estimates (Wei and Simko, 2017), and factor analysis results are provided for all six tools.

Analyses for all tools are presented simultaneously in the following order:

- MVN assumption assessment
- tile plots
- correlation coefficient estimates and visualizations

- graphical display of results of factor analysis by minres

3.1.6.1 Variable reduction in MDOT, BLM and NRCS tools To conduct factor analysis, Tools with a large number of metrics relative to the sample size necessitated variable reduction. Due to the limited sample size (n=18), the data were approaching square and needed to undergo variable reduction in order for correlation coefficients to be estimated, and factor analysis to be a reasonable approach. The MDOT, BLM, and NRCS tools had metrics combined and/or removed in a logical fashion based upon expert opinion. Although variable reduction is not ideal, this process represents a common situation – wetland evaluation is time-consuming and expensive; thus the sample size is often small. Tools relying on a large number of indices can then be difficult to assess and interpret. If necessary, the process of variable reduction will vary depending on the sample, sample size, and tools under consideration. Different variable reduction processes for a single tool may lead to different results. Managers merging indices for various tools should aim to have no fewer than four re-grouped indices. Any fewer number of indices could result in potentially unreliable results.

The NRCS tool is comprised of ten metric indices. The categories, *stream incisement*, *streambanks with excessive lateral cutting*, and *stream in balance with water and sediment supplied by watershed* were merged into a single index now labeled *geomorphic*. The categories *streambank with vegetation having a deep, binding root mass*, *riparian/wetland vegetation cover in the riparian/flood plain area*, *noxious weeds in the riparian area*, *introduced/exotic undesirable plants in the riparian area* and *woody species establishment and regeneration* were merged into a single index now labeled *vegetation*. The original categories *flood plain characteristics for dissipating energy and capturing sediment* and *utilization of trees and shrubs* were retained, yielding four total metric categories. Overall scores were re-averaged based upon the new condensed metrics.

The MWAM is comprised of 12 metric indices. The category *habitat for federally listed or proposed threatened or endangered (T&E) plants or animals* was removed because scores for all sites were zero. *General wildlife habitat* and *general fish habitat* were merged into a single index now labeled *general habitat*. The indices *sediment/nutrient/toxicant retention and removal* and *sediment/shoreline stabilization* were merged into a single index now labeled *sediment source/storage*. The indices *short and long term surface water storage* and *groundwater discharge/recharge* were merged into a single index labeled *water storage/exchange*. *Habitat for plants or animals rated S1, S2, or S3 by the MTNHP*, *flood attenuation*, and *production export/terrestrial and aquatic food chain support* were retained as is. The index *uniqueness* was retained. These changes yield seven total metric categories. Overall scores were re-averaged based upon the new condensed metrics.

The PFC is comprised of seventeen metric indices. The Riparian area management manual (USDI, 2015) presented these seventeen indices in three general groups, and metrics were re-merged according to these groupings: *hydrology*, *vegetation* and *geomorphology*. Although it would be ideal to use these three indices, such a small number of indices may lead to unreliable results. Therefore, the PFC was condensed to four indices: *hydrology*, *vegetation*, *geomorphology*, and *carbon*. Overall scores were re-averaged based upon the new condensed metrics.

3.1.6.2 Tile plots, initial data visualization Tile plots are a useful way of visualizing high-dimensional data. Figure 9 displays tile plots for each tool, allowing us to visually compare scoring patterns within each index over the range of sites, and how similar these patterns are to the overall site score pattern. Multi-metric indices and the overall score by site appear on the y-axis. Score is on the x-axis, arranged in descending order of overall site score from left to right. Coloring indicates score, with lower scores in lighter colors while darker colors indicate higher scores. Recall that index scores are eventually averaged to create the overall site score which describes the latent variable 'ecological health'.

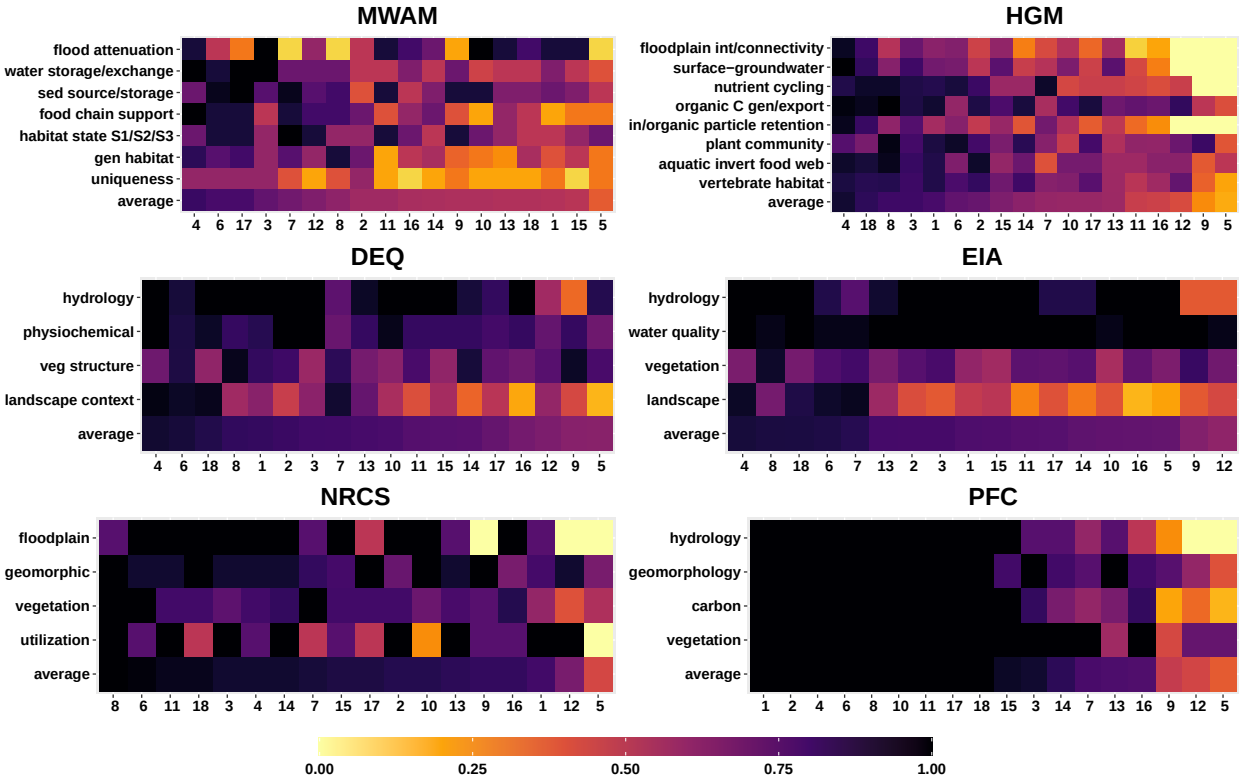


Figure 9: Tile plot of index scores and overall tool score by site for each tool. Indices and overall average score are listed on the y-axes, with sites on the x-axes. Moving from the left to the right-hand side of the x-axis, sites are arranged in descending order of overall tool score. Coloring indicates score, with lower scores in lighter colors and higher scores in darker colors.

3.1.6.3 Correlation coefficient estimates The usual, more formal, measure of similarity or association between variables is the correlation coefficient. Figure 10 shows correlation matrices for the indices of each tool. The color legend below the plot shows dark blues representing strongly positive linear relationships, dark reds strongly negative linear relationships, and lighter colors tending towards white indicating no linear relationship. Shape also indicates strength of relationships with smaller circles indicating a lack of a relationship and larger circles indicating a stronger linear relationship.

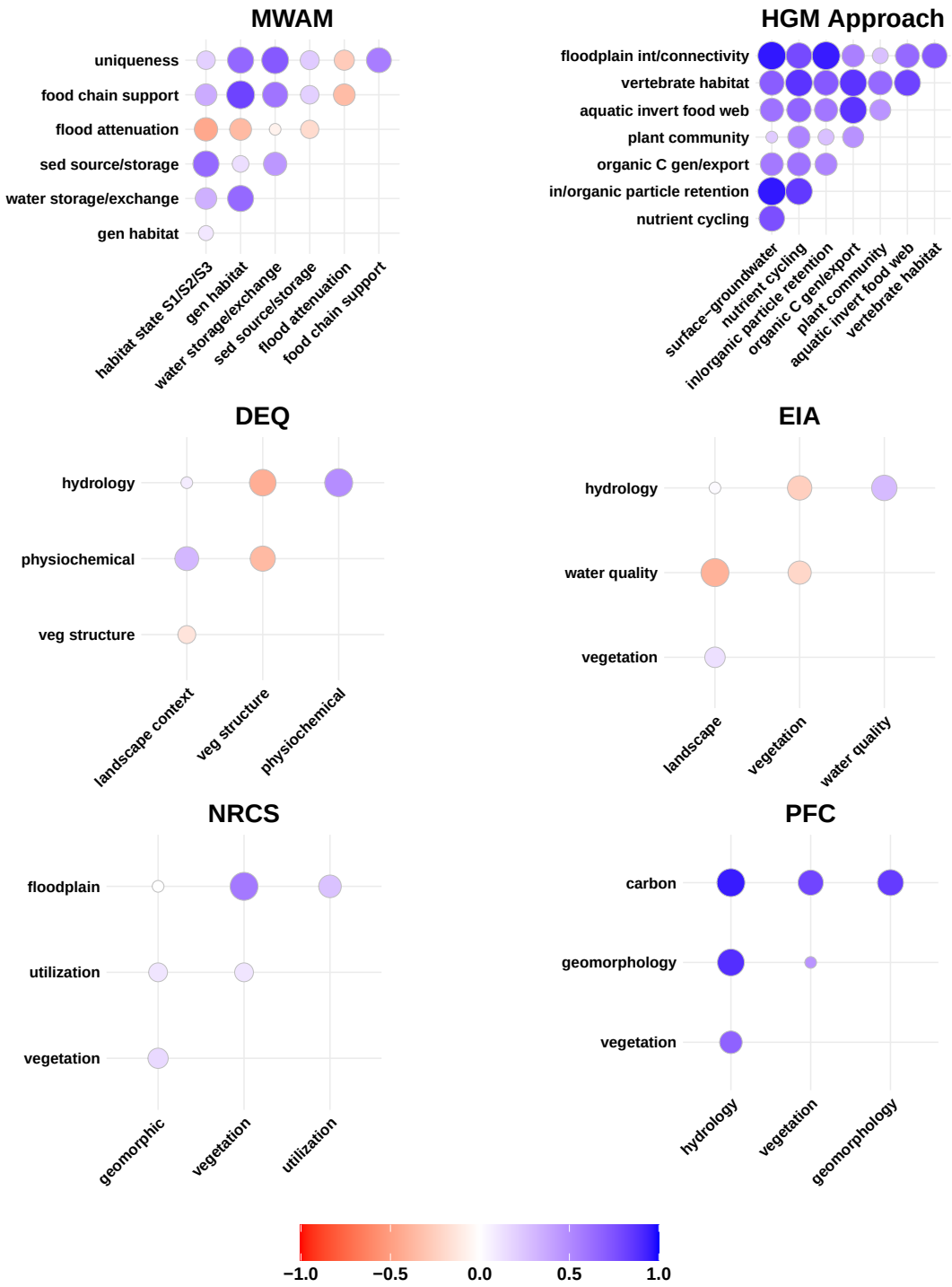


Figure 10: Correlation matrices for the indices of each tool. The color legend below the plot shows dark blues representing strongly positive linear relationships, dark reds strongly negative linear relationships, and lighter colors tending towards white indicating no linear relationship. Shape also indicates strength of relationships with smaller circles indicating a lack of a relationship and larger circles indicating a stronger linear relationship.

3.1.6.4 Maximum likelihood appraoch to EFA using least squares minimum residuals The 'fa' function in the *psych* package (Revelle, 2019) uses one of six different algorithms: minimum residual (minres, aka ols, uls), principal axes, alpha factoring, weighted least squares, minimum rank, or maximum likelihood. "The minimum residual solution is an unweighted least squares solution that takes a slightly different approach. It uses the optim function and adjusts the diagonal elements of the correlation matrix to minimize the squared residual when the factor model is the eigenvalue decomposition of the reduced matrix." Here we have chosen to use the minimum residuals algorithm because "this produces solutions very similar to maximum likelihood even for badly behaved matrices... (it) will work when maximum likelihood will not for (it) can be used when the matrix is singular" (Revelle, 2019). This is especially useful for our situation, which is complicated by a small sample size.

Results of factor analysis by minres are displayed for each tool in the form of a bar chart and numerical estimates of factor loadings (Figure 11). Factor loadings are "a measure of the degree of generalizability found between each variable and (a) factor... reflect(ing) quantitative relationships. The farther a factor loading is from zero, the more one can generalize from that factor to the variable" (Gorsuch, 1983). Here, as we assume a single factor – ecological health – exists, we are interested in exploring how generalizable the indices of each tool are to the overall average tool score that describes ecological health. Weak or negative factor loadings suggest that some of the individual health indicators/metrics are not contributing to the concept of 'ecological health' in the way that the tool assumes.

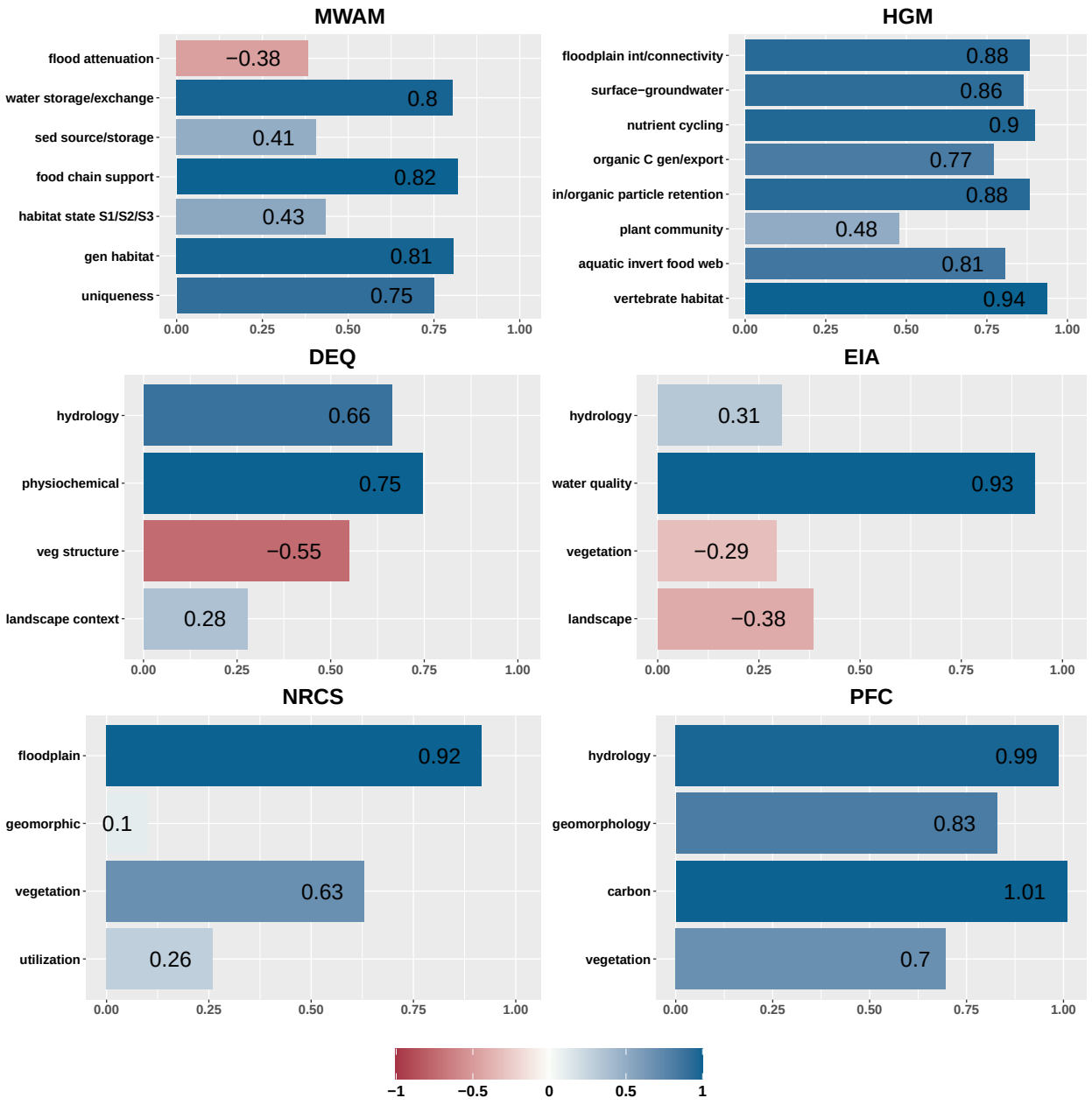


Figure 11: Results of factor analysis for each tool displayed in the form of a bar chart and numerical estimates of factor loadings. The absolute value of estimated factor loadings is plotted for an easier comparison between index loadings for each tool. Coloring indicates negative or positive loadings: loadings approaching -1 are darker red, loadings near zero tend to white, and loadings approaching 1 are blue. The legend appears at the bottom of the figure. Individual metrics for each tool are listed on the y-axis.

4 Discussion

4.1 Interpretation of results

Factor analysis results for NRCS estimate a strong loading only for 'floodplain dissipates energy' (0.92), and a moderately strong loading for 'vegetation' (0.63). 'Utilization of trees/shrubs' has a weak loading (0.26) and the index 'geomorphic' has a very weak estimated factor loading of 0.10. These mixed results indicate that not all indices generalize well to 'ecological health', and that for this sample, the model may not be functioning as assumed.

MWAM results show negative, moderate, and strong loadings. 'Flood attenuation' has a negative estimated factor loading (-0.38), 'habitat state S1/S2/S3' (0.43) and 'sediment source/storage' (0.41) have weaker loadings, 'uniqueness' has a moderate to strong estimated loading (0.75) and the remaining indices all have strong loadings: 'water storage/exchange' (0.80), 'general habitat' (0.81) and food chain support (0.82). These results suggest that not although the majority of indices generalize well to a single factor, this is not the case for all the factors - particularly 'flood attenuation' - and the model may not be functioning as assumed.

The HGM Approach shows seven of the eight indices with strong estimated loadings onto 'ecological health' ranging from 0.77-0.94: 'vertebrate habitat' (0.94), 'nutrient cycling' (0.90), 'in/organic particle retention' (0.88), 'floodplain in/connectivity' (0.88), 'surface-groundwater' (0.86), 'aquatic invertebrate food web' (0.81), and 'organic C gen/export' (0.77). Only the index 'plant community' (0.48) has a factor loading falling below 0.5. However, it is important to note that the HGM uses 13 actual measurements to create eight indices. The indices are not direct measurements but rather differently scaled averages of various combinations of these 13 measurements (US Army Corps of Engineers, 2002). Many of the indices utilize the same underlying measurements in their calculations - and will therefore be related by this fact alone - so these results should be interpreted cautiously.

The DEQ and EIA, despite their common foundations, behave differently. For the DEQ, 'phyisochemical' has the strongest loading of 0.75, followed by 'hydrology' with a moderate loading (0.66). 'Landscape context' has a weak estimated loading (0.28) and 'vegetation structure' has a negative loading (-0.55). Two of the four indices do not generalize well to 'ecological health', suggesting that for this sample this tool may not behave as expected. EIA results have a strong estimated loading only for the 'water quality' index (0.93). Otherwise, 'hydrology' has a weak loading (0.31), and both 'vegetation' and 'landscape' have negative estimated loadings (-0.29 and -0.38, respectively). For this sample, three of the four EIA indices do not generalize well to 'ecological health'.

Results of factor analysis for the PFC are problematic. 'Hydrology' (0.99) and 'geomorphology' (0.834) both load strongly onto 'ecological health' and 'vegetation' has a moderate to strong loading (0.70). The estimated loading for 'carbon' is 1.01 and would be considered an ultra-Heywood case, rendering the factor solution invalid (SAS manual). This estimate is on the cusp between a Heywood case (loading estimate of 1) and an ultra-Heywood case, although disagreement about the validity of a Heywood case exists (SAS manual). Both should be considered suspicious results, as they imply either a negative estimate of the variable's specific variance (ultra-Heywood case) or a specific variance estimate of zero (Heywood case), which is unlikely in practice (Everitt & Hothorn, 2011). In our case, these results may be due to bad prior communality estimates, too few common factors, a small sample size, or that the common factor model is not an appropriate model for the data.

As a test, the number of factors was increased to two, and although no Heywood cases were observed, the results were not easily interpretable, suggesting that this may not be the main issue. We should also reiterate that this analysis is based upon one possible reorganization of the original 17 indices into four indices and that results may be different given a different variable reduction process. We already recognize that our sample size is small, and a larger sample would help us to better understand if the common factor model is appropriate for these data.

4.2 Limitations

Sample size is the most glaring limitation of this study. The number of original health indicator categories for the MWAM, PFC and NRCS tools needed to be reduced and/or condensed due to the number of sites under consideration (n=18). As with any other analysis, a large enough sample size can yield more reliable results. But what is 'large enough'? A subject to item ratio of 10:1 is a rule of thumb still in use for exploratory factor analyses. According to Casello and Osborne, 'strong data' as defined by "uniformly high communalities without cross loadings, plus several variables loading strongly on each factor" may allow a smaller sample size. Thus, the important issues under consideration for reliable results include the number of items loading on each factor, cross-loadings (i.e. items loading on more than one factor), and items with low communalities which may suggest an additional factor that should be explored.

Because our study is only interested in assessing the efficiency of each tool in distinguishing between high and low-scoring sites, we have avoided these issues. Tool developers have already decided that only one factor exists, and have therefore determined it sufficient to take an average of all health indicators to produce a single summary statistic that describes the 'ecological health' of a given site, removing issues with cross-loadings and pre-determining the appropriate number of factors. Results that indicate low communalities for an item or items only suggest that taking an average of the health indicators for that tool may not be a useful way of summarizing the data at hand. This is what we seek to assess - the strengths and weaknesses of each tool and why this may be the case, as opposed to determining the true underlying structure of the data. In this sense, the sample size under consideration does not prevent this analysis from revealing useful information.

5 Citations

- Auguie, B. (2017). gridExtra: Miscellaneous Functions for "Grid" Graphics. R package version 2.3. https://CRAN.R-project.org/package=gridExtra
- Berglund, J. and McEldowney, R. 2008. MDT Montana wetland assessment method. Prepared for: Montana Department of Transportation. Post, Buckley, Schuh & Jernigan. Helena, Montana. 42 p.
- Costello, A. and Osborne, J. (2005). Best Practices in Exploratory Factor Analysis: Four Recommendations for Getting the Most From Your Analysis. *Practical Assessment, Research & Evaluation.* 10(7). https://doi.org/10.7275/jyj1-4868
- Daróczi, G. and Tsegelskyi, R. (2018). pander: An R 'Pandoc' Writer. R package version 0.6.3. https://CRAN.R-project.org/package=pander
- Everitt, B. and Hothorn, T. (2011). *An Introduction to Applied Multivariate Analysis with R.* Springer Science and Business Media, LLC. DOI 10.1007/978-1-4419-9650-3_5.
- Garnier, S. (2018). viridis: Default Color Maps from 'matplotlib'. R package version 0.5.1. https://CRAN.R-project.org/package=viridis
- Gorsuch, R. (1983). *Factor Analysis* (2nd ed.).Lawrence Erlbaum Associates, Inc.
- Joreskog, K. (2003). *Factor Analysis by MINRES.* https://pdfs.semanticscholar.org/5aa4/ff22f6aae72bc52548b3ab59bc2a6852d18d.pdf
- Kassambara, A. (2020). ggpubr: 'ggplot2' Based Publication Ready Plots. R package version 0.4.0. https://CRAN.R-project.org/package=ggpubr
- Montana Department of Environmental Quality. (2019). *Montana DEQ Wetland Program Wetland Assessment Protocol - Field Manual.* URL
- Montana Natural Heritage Program. (2018). *Montana Ecological Integrity Assessment Field Manual.* URL
- Neuwirth, E. (2014). RColorBrewer: ColorBrewer Palettes. R package version 1.1-2. https://CRAN.R-project.org/package=RColorBrewer
- R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.
- Revelle, W. (2019). psych: Procedures for Personality and Psychological Research, Northwestern University, Evanston, Illinois, USA, https://CRAN.R-project.org/package=psych Version = 1.9.12.
- United States Army Corps of Engineers. (2002). *A Regional Guidebook for Applying the Hydrogeomorphic Approach to Assessing Wetland Functions of Riverine Floodplains in the Northern Rocky Mountains* http://www.wes.army.mil/el/wetlands/wlpubs.html
- U.S. Department of the Interior, 2015. Riparian area management: Proper functioning condition assessment for lotic areas. Technical Reference 1737-15. Bureau of Land Management, National Operations Center, Denver, CO. http://www.blm.gov/wo/st/en/infolblm-library/publications/blm_publications/tech_refs.html
- United States Department of Agriculture, Natural Resources Conservation Service. (2012). *Riparian Assessment Environmental Technical Note No. MT-2 (Rev. 1)* http://www.mt.nrcs.usda.gov/technical/
- Urbanek, S. (2019). jpeg: Read and write JPEG images. R package version 0.1-8.1. https://CRAN.R-project.org/package=jpeg

- Wei, T. and Simko, V. (2017). R package "corrplot": Visualization of a Correlation Matrix (Version 0.84). Available from https://github.com/taiyun/corrplot
- Wickham, H. and Bryan, J. (2019). readxl: Read Excel Files. R package version 1.3.1. https://CRAN.R-project.org/package=readxl
- Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T., Miller, E., Bache, S., Müller, K., Ooms, J., Robinson, D., Seidel, D., Spinu, V., Takahashi, K., Vaughan, D., Wilke, C., Woo, K., and Yutani, H. (2019). Welcome to the tidyverse. Journal of Open Source Software, 4(43), 1686, https://doi.org/10.21105/joss.01686

6 Code

```
knitr::opts_chunk$set(echo = FALSE, fig_caption = TRUE,
                      message = FALSE, warning=FALSE)
#set working directory for entire document instead of a single chunk
knitr::opts_knit$set(root.dir="C:/Users/R/Desktop/assessment_review/")
require(pacman)
pacman::p_load(tidyr, tidyverse, knitr, readr, pander, tinytex, readxl,
               RColorBrewer, ggpubr, cowplot, gridExtra, jpeg, grid, ggcorrplot,
               psych, viridis, skimr, devtools, Rmisc, ggrepel, bookdown)

p2 <- readxl::read_excel("assessment_121620.xlsx",
                         sheet = "Normalized", na = "9999") %>%
  dplyr::rename(atool = `...1`, indicator = Wetland) %>%
  dplyr::mutate(indicator = ifelse((indicator == "Average" |
                                      indicator == "Total site Score" |
                                      indicator == "Whole score" |
                                      indicator == "Final Score" |
                                      indicator == "Overall Rating" |
                                      indicator == "Properly Funcitoning Condition"),
                                   "average", indicator),
                atool = factor(atool, levels=c("BLM", "DEQ", "HGM", "MDOT",
                                               "MTNHP", "NRCS")),
                indicator = ifelse(indicator=="streambanks are laterally stable",
                                   "streambanks lat stable",
                           ifelse(indicator=="stream system is vertically stable",
                                   "stream sys vert stable",
                           ifelse(indicator=="stream is in balance with the water/sediment",
                                   "stream balanced with water/sed",
                           ifelse(indicator=="stabilizing communities capable of
                                   withstanding highwater",
                                   "stab comm withstand high water",
                           ifelse(indicator=="species indicate maintenance of
                                   soil moisture",
                                   "species maintain soil moisture",
                           ifelse(indicator=="Sinuosity, gradient, w/d ratio are balanced",
                                   "sin/grad/w-d ratio balanced",
                           ifelse(indicator=="Riparian plants exhibit high vigor",
                                   "plants exhibit vigor",
                           ifelse(indicator=="Riparian impairment is absent",
                                   "impairment absent",
                           ifelse(indicator=="Riparian area is expanding or max
                                   potent extent",
                                   "area expanding/max extent",
                           ifelse(indicator=="Point bars are reveg with stabilizing plants",
                                   "point bars revegetated",
                           ifelse(indicator=="Plant communities are an adequate
                                   source of  woody materials",
                                   "plants adeq woody source",
                           ifelse(indicator=="Floodplain/Channel characteristics
                                   dissipate energy",
                                   "floodplain dissipates energy",
                           ifelse(indicator=="Floodplain inundated frequently",
```

```
                              "floodplain inundated freq",
                       ifelse(indicator=="Diverse stabilizing veg for
                              recover/maintenance",
                              "diverse veg recovery/maintenance",
                       ifelse(indicator=="Beaver dams are stable",
                              "beaver dams stable",
                       ifelse(indicator=="An adequate stabling riparian
                              veg protect banks/dissipate energy",
                              "veg protects bank/dissipates energy",
                       ifelse(indicator=="adequate age classes of stabilizing
                              veg recover/maintain",
                              "age veg recovery/maintenance", indicator)))))))))))))))))),
    #rename DEQ indicators
    indicator = ifelse(indicator=="Landscape context",
                       "landscape context",
                       ifelse(indicator=="Vegetation structure AA",
                              "veg structure",
                       ifelse(indicator=="physiochemical metrics  AA",
                              "physiochemical",
                       ifelse(indicator=="Hydrology Indicators",
                              "hydrology", indicator)))),
    #rename MTNHP indicators
    indicator = ifelse(indicator=="Landscape Metrics",
                       "landscape",
                       ifelse(indicator=="Vegetation Metrics In the AA",
                              "vegetation",
                       ifelse(indicator=="Water Quality",
                              "water quality",
                       ifelse(indicator=="Hydrology",
                              "hydrology", indicator)))),
    #rename MDOT indicators
    indicator = ifelse(indicator=="hab_Fed_rate",
                       "habitat fed T&E ",
                       ifelse(indicator=="hab_state_rate",
                              "habitat state S1/S2/S3",
                       ifelse(indicator=="hab_ gen_rate",
                              "gen wildlife habitat",
                       ifelse(indicator=="hab_fish_rate",
                              "gen fish habitat",
                       ifelse(indicator=="flood_attn",
                              "flood attenuation",
                       ifelse(indicator=="wtr_stor",
                              "surface water storage",
                       ifelse(indicator=="sed_stor",
                              "sed/nut/tox retention/removal",
                       ifelse(indicator=="bank_stab",
                              "sed/shore stabilization",
                       ifelse(indicator=="food_chain_suport",
                              "food chain support",
                       ifelse(indicator=="GW recharge rating",
                              "groundwater dis/recharge",
                       ifelse(indicator=="Uniqueness",
                              "uniqueness", indicator)))))))))))),
```

```
        #rename HGM indicators
        indicator = ifelse(indicator=="Surface-groundwater",
                           "surface-groundwater",
                           ifelse(indicator=="Nutrient Cycling",
                                  "nutrient cycling",
                           ifelse(indicator=="Retentionof Organic and Inorganic partiles",
                                  "in/organic particle retention",
                           ifelse(indicator=="Gen and export of organic carbon",
                                  "organic C gen/export",
                           ifelse(indicator=="characteristic plant community",
                                  "plant community",
                           ifelse(indicator=="Char Aquatic Invert food web",
                                  "aquatic invert food web",
                           ifelse(indicator=="char vert habitat",
                                  "vertebrate habitat",
                           ifelse(indicator=="Floodplain interspesion",
                                  "floodplain int/connectivity",
                                  indicator)))))))),
        #rename NRCS indicators
        indicator = ifelse(indicator=="Stream Incisement",
                           "stream incisement",
                           ifelse(indicator=="Streambank stability",
                                  "streambank stability",
                           ifelse(indicator=="Sediment Transport",
                                  "sediment transport",
                           ifelse(indicator=="Streambank Root Mass",
                                  "streambank root mass",
                           ifelse(indicator=="Noxious Week Cover",
                                  "noxious weeds",
                           ifelse(indicator=="Riparain Veg Cover",
                                  "floodplain veg cover",
                           ifelse(indicator=="Distrubed Veg Composition",
                                  "introduced/exotic plants",
                           ifelse(indicator=="Woody Regeneration",
                                  "woody species regeneration",
                           ifelse(indicator=="Woody Browsing",
                                  "utilization trees/shrubs",
                           ifelse(indicator=="Floodplain roughness",
                                  "floodplain dissipates energy",
                                  indicator))))))))))) %>%
  dplyr::filter(!(indicator == "Total" |
                    indicator == "MAX" |
                    indicator == "Impact of disturbances" |
                    indicator == "Disturbance Scope rating" |
                    indicator == "Impact to Function")) %>%
  dplyr::mutate(atool= ifelse(atool=="MDOT", "MWAM",
                              ifelse(atool=="MTNHP", "EIA",
                                     ifelse(atool=="BLM", "PFC",
                                            as.character(atool)))))

#push from wide to long format
p_long <- p2 %>% tidyr::gather(key="site", value="score", c(3:20)) %>%
  #rename BLM indicators
```

```
  dplyr::mutate(score=as.numeric(as.character(score)))

final <- p2 %>% filter(indicator == "New LDI") %>%
  tidyr::gather(key="site", value="score", c(3:20)) %>%
  dplyr::mutate(grouping=as.numeric(as.character(score)),
                grouping=ifelse(grouping < 2.6, "mainly intact",
                                ifelse(grouping >= 2.6 & grouping <= 4.5,
                                       "mid-range", "most disturbed") )) %>%
  select(., grouping, site) %>%
  right_join(., y=p_long, by="site")
final %>%
  mutate(atool = ifelse(indicator=="New LDI", "LDI",
                        as.character(atool)),
         atoolf = factor(atool,
                         levels = c("LDI", "MWAM","PFC","HGM",
                                    "DEQ", "EIA","NRCS")),
         score = ifelse(indicator == "New LDI", 1-(score/8.97), score)) %>%
  filter(indicator=="average" | indicator == "New LDI") %>%
  select(., atoolf, score) %>% group_by(atoolf) %>%
  skimr::skim() %>%
  dplyr::select(atoolf, numeric.mean, numeric.p0, numeric.p100) %>%
  mutate(`assessment tool`=atoolf, `mean average score`=round(numeric.mean, 2),
         `min score` = round(numeric.p0, 2),
         `max score`=round(numeric.p100, 2)) %>%
  dplyr::select(`assessment tool`,
                `mean average score`,
                `min score`, `max score`) %>%
  mutate(`score range` = round(`max score`- `min score`, 2)) %>%
  pander(caption = "Summary statistics across sites by assessment tool. Average overall score, minimum

final %>%
  mutate(atool = ifelse(indicator=="New LDI",
                        "LDI",
                        as.character(atool)),
         atoolf = factor(atool, levels = c("PFC", "NRCS",
                                           "EIA", "DEQ", "HGM", "MWAM", "LDI")),
         score = ifelse(indicator == "New LDI", 1-(score/8.97), score)) %>%
  filter(indicator=="average" | indicator == "New LDI") %>%
  ggplot2::ggplot(ggplot2::aes(y=atoolf, x=score, color=grouping,
                               shape=grouping)) +
  #add a point of a slightly bigger size to create a black outline look
  geom_point(position=position_dodge(width=0.4), size=5) +
  theme_bw() +
  ggplot2::theme(axis.text.y = ggplot2::element_text(angle=0, size=13,
                                                     face="bold", color="black"),
                 axis.text.x = ggplot2::element_text(angle=0, size=13,
                                                     face="bold", color="black"),
                 legend.title = element_text(size=13, face="bold"),
                 legend.text = element_text(size=12),
                 legend.position = c(0.13, 0.26),
                 legend.background = element_rect(fill="lightgrey",
                                  size=1, linetype="solid",
                                  colour ="black"),
```

```
                legend.key.size =  unit(1, "cm"),
                legend.box="horizontal",
                axis.title.x=element_text(size=14, face="bold"),
                plot.title=element_text(size=22),
                axis.ticks = element_blank(),
                panel.grid.major = element_line(colour="grey70", size=0.3)) +
                #panel.grid.minor = element_blank()) +
  ggplot2::scale_x_continuous(limits = c(0, 1), breaks=seq(0, 1, 0.1)) +
  ggplot2::labs(title="Overall Site Scores by Assessment Tool",
                x="overall site score",
                y="") +
  scale_colour_manual(name = "LDI grouping",
                      values = c("#009E73", "#F0E442", "#000000"),
                      labels = c("mainly intact", "mid-range", "most disturbed")) +
  scale_shape_manual(name="LDI grouping",
                     values = c("triangle", "square", "circle"),
                     labels = c("mainly intact", "mid-range", "most disturbed"))
#display high-scoring sites via table
final %>%
  mutate(atool = ifelse(indicator=="New LDI",
                        "LDI", as.character(atool)),
         atoolf = factor(atool, levels = c("PFC", "NRCS",
                                           "EIA", "DEQ", "HGM",
                                           "MWAM", "LDI")),
         score = ifelse(indicator == "New LDI", 1-(score/8.97), score)) %>%
  filter(grouping=="mainly intact" & (indicator=="average" |
                                        indicator=="New LDI")) %>%
  select(., site, atoolf, score) %>%
  group_by(site, atoolf, score) %>%
  dplyr::rename(`assessment tool`=atoolf) %>%
  pander(caption="Mostly intact sites (4, 6, 7, 8, and 18) overall site scores by assessment tool. LDI s
final %>%
  filter(grouping=="mainly intact" & (indicator=="average" |
                                        indicator == "New LDI")) %>%
  mutate(atool = ifelse(indicator=="New LDI", "LDI",
                        as.character(atool)),
         atoolf = factor(atool,
                         levels = c("LDI", "MWAM", "HGM", "DEQ",
                                    "EIA","NRCS", "PFC"))) %>%
  select(., atoolf, score) %>%
  group_by(atoolf) %>%
  skimr::skim() %>%
  dplyr::select(atoolf, numeric.mean, numeric.p0, numeric.p100) %>%
  mutate(`assessment tool`=atoolf,
         `mean average score`=round(numeric.mean, 2),
         `min score` = round(numeric.p0, 2),
         `max score`=round(numeric.p100, 2)) %>%
  dplyr::select(`assessment tool`, `mean average score`,
                `min score`, `max score`) %>%
  mutate(`score range` = round(`max score`- `min score`, 2)) %>%
  pander(caption = "Summary statistics across mainly intact sites by assessment tool. Average overall s
#plot mostly intact sites by score and tool
#create vertical line data with LDI scores for each site
```

```
LDIgroup1 <- p2 %>% dplyr::filter(indicator=="New LDI") %>%
  tidyr::gather(key="site", value="score", c(3:20)) %>%
  mutate(score=1-(score/8.97)) %>% dplyr::select(-atool) %>%
  filter(site=="Site 4" | site == "Site 6" | site=="Site 7" |
           site=="Site 8" |site=="Site 18") %>%
  mutate(sitef = factor(site, levels=c("Site 7", "Site 8",
                                       "Site 18", "Site 4", "Site 6")))

p <- final %>%
  filter(grouping=="mainly intact" & indicator=="average") %>%
  mutate(atoolf = factor(atool, levels = c("PFC", "NRCS", "EIA",
                                           "DEQ", "HGM", "MWAM", "LDI"))) %>%
  #change site numbers to be ordered by LDI values in faceting
  mutate(sitef = factor(site, levels=c("Site 7", "Site 8",
                                       "Site 18", "Site 4", "Site 6"))) %>%
  ggplot2::ggplot(ggplot2::aes(y=atoolf, x=score, color=atoolf)) +
  ggplot2::geom_point(ggplot2::aes(y=atoolf, x=score, color=atoolf), size=6) +
  ggplot2::theme(axis.text.y = ggplot2::element_text(angle=0, size=12,
                                                     face="bold", color="black"),
                 axis.text.x = ggplot2::element_text(angle=0, size=12,
                                                     face="bold", color="black"),
                 strip.text.y = element_text(size = 10,
                                             colour = "black", angle = 90),
                 legend.position = "none",
                 axis.title.x=element_text(size=14, face="bold"),
                 plot.title=element_text(size=22)) +
  ggplot2::scale_x_continuous(limits = c(0, 1), breaks=seq(0, 1, 0.1)) +
  ggplot2::facet_grid(sitef~.) +
  ggplot2::labs(title="Mainly Intact Sites (7, 8, 18, 4, 6): Overall Score by Assessment Tool",
                x="overall (average) score",
                y="",
                color="assessment tool")
#use this code if LDI are rescaled to the opposite direction and on a scale 0-1
p + ggplot2::geom_vline(ggplot2::aes(xintercept = score), LDIgroup1,
                        linetype="solid", color="black", size=2)


#create vertical line data with averages for each tool
averages <- p_long %>% dplyr::filter(indicator=="average") %>%
  dplyr::select(c("atool", "score", "site"))
#select for specific sites for various graphics
vline.data1 <- averages %>% dplyr::filter(site=="Site 7") %>%
  dplyr::select(c("atool", "score"))
#create color palette
cols <- RColorBrewer::brewer.pal(n=6, "Dark2")

plot1 <- p_long %>%
  dplyr::filter(., site == 'Site 7' &
                  !(indicator=="average") &
                  !(indicator=="New LDI")) %>%
  mutate(atoolf = factor(atool, levels = c("PFC", "NRCS",
                                           "EIA", "DEQ", "HGM",
                                           "MWAM", "LDI"))) %>%
```

```
  ggplot2::ggplot(ggplot2::aes(y=indicator, x=score, color=atoolf)) +
  ggplot2::geom_point(ggplot2::aes(y=indicator, x=score, color=atoolf), size=4) +
  ggplot2::theme(axis.text.y = ggplot2::element_text(angle=0),
                 legend.position = "none",
                 axis.title=element_text(size=12, face="bold"),
                 plot.title=element_text(size=18),
                 plot.subtitle = element_text(size=14),
                 strip.text = element_text(size=10, face="bold")) +
  ggplot2::scale_x_continuous(limits = c(0, 1), breaks=seq(0, 1, 0.2)) +
  ggplot2::facet_grid(atool~., scales="free_y", space="free") +
  ggplot2::labs(title="Site 7: Indicator and Average Comparisons",
                subtitle="Faceted by assessment tool",
                x="indicator score",
                y="",
                color="assessment tool") +
ggplot2::scale_colour_manual(values=cols)

plot1 + ggplot2::geom_vline(ggplot2::aes(xintercept = score),
                            vline.data1, linetype="solid",
                            color="black", size=2)

#summarize mid-range sites via table
final %>%
  mutate(atool = ifelse(indicator=="New LDI", "LDI", as.character(atool)),
         atoolf = factor(atool,
                         levels = c("PFC", "NRCS", "EIA",
                                    "DEQ", "HGM", "MWAM", "LDI")),
         score = ifelse(indicator == "New LDI", 1-(score/8.97), score)) %>%
  filter(grouping=="mid-range" & (indicator=="average" | indicator=="New LDI")) %>%
  select(., site, atoolf, score) %>%
  group_by(site, atoolf, score) %>%
  dplyr::rename(`assessment tool`=atoolf) %>%
  pander(caption="Mid-range sites (3, 9, 10, 12, 13, and 17) overall site scores per assessment tool. LI
final %>%
  filter(grouping=="mid-range" & (indicator=="average" |
                                    indicator == "New LDI")) %>%
  mutate(atool = ifelse(indicator=="New LDI", "LDI",
                        as.character(atool)),
         atoolf = factor(atool, levels = c("LDI", "MWAM", "HGM",
                                           "DEQ", "EIA","NRCS", "PFC"))) %>%
  select(., atoolf, score) %>%
  group_by(atoolf) %>%
  skimr::skim() %>%
  dplyr::select(atoolf, numeric.mean, numeric.p0, numeric.p100) %>%
  mutate(`assessment tool`=atoolf,
         `mean average score`=round(numeric.mean, 2),
         `min score` = round(numeric.p0, 2),
         `max score`=round(numeric.p100, 2)) %>%
  dplyr::select(`assessment tool`, `mean average score`,
                `min score`, `max score`) %>%
  mutate(`score range` = round(`max score`- `min score`, 2)) %>%
  pander(caption = "Summary statistics across mid-range sites by assessment tool. Average overall score
#plot mid-range sites by score and tool
```

```
#create vertical line data with LDI scores for each site
LDIgroup2 <- p2 %>% dplyr::filter(indicator=="New LDI") %>%
  tidyr::gather(key="site", value="score", c(3:20)) %>%
  mutate(score=1-(score/8.97)) %>% dplyr::select(-atool) %>%
  filter(site=="Site 3" | site == "Site 9" | site=="Site 10" |
           site=="Site 12" |site=="Site 13" | site=="Site 17") %>%
  mutate(sitef = factor(site, levels=c("Site 13", "Site 17",
                                       "Site 9", "Site 12",
                                       "Site 3","Site 10")))

p <- final %>%
  filter(grouping=="mid-range" & indicator=="average") %>%
  mutate(atoolf = factor(atool,
                         levels = c("PFC", "NRCS", "EIA",
                                    "DEQ", "HGM", "MWAM", "LDI"))) %>%
  mutate(sitef = factor(site,
                        levels=c("Site 13", "Site 17", "Site 9",
                                 "Site 12", "Site 3","Site 10"))) %>%
  ggplot2::ggplot(ggplot2::aes(y=atoolf, x=score, color=atoolf)) +
  ggplot2::geom_point(ggplot2::aes(y=atoolf, x=score, color=atoolf), size=6) +
  ggplot2::theme(axis.text.y = ggplot2::element_text(angle=0, size=12,
                                                     face="bold", color="black"),
                 axis.text.x = ggplot2::element_text(angle=0, size=12,
                                                     face="bold", color="black"),
                 strip.text.y = element_text(size = 10,
                                             colour = "black", angle = 90),
                 legend.position = "none",
                 axis.title.x=element_text(size=14, face="bold"),
                 plot.title=element_text(size=22)) +
  ggplot2::scale_x_continuous(limits = c(0, 1), breaks=seq(0, 1, 0.1)) +
  ggplot2::facet_grid(sitef~.) +
  ggplot2::labs(title="Mid-range Sites (10, 3, 12, 9, 17, 13): Overall Scores by Assessment Tool",
                x="overall (average) score",
                y="",
                color="assessment tool")
#use this code if LDI are rescaled to the opposite direction and on a scale 0-1
p + ggplot2::geom_vline(ggplot2::aes(xintercept = score),
                        LDIgroup2, linetype="solid", color="black", size=2)

#plot site 3 comparisons

#create vertical line data with averages for each tool
averages <- p_long %>% dplyr::filter(indicator=="average") %>%
  dplyr::select(c("atool", "score", "site"))
#select for specific sites for various graphics
vline.data1 <- averages %>% dplyr::filter(site=="Site 13") %>%
  dplyr::select(c("atool", "score"))
#create color palette
cols <- RColorBrewer::brewer.pal(n=6, "Dark2")

plot1 <- p_long %>% dplyr::filter(., site == 'Site 13' &
                                    !(indicator=="average") &
                                    !(indicator=="New LDI")) %>%
```

```
  mutate(atoolf = factor(atool,
                         levels = c("PFC", "NRCS", "EIA",
                                    "DEQ", "HGM", "MWAM", "LDI"))) %>%
  ggplot2::ggplot(ggplot2::aes(y=indicator, x=score, color=atoolf)) +
  ggplot2::geom_point(ggplot2::aes(y=indicator, x=score, color=atoolf), size=4) +
  ggplot2::theme(axis.text.y = ggplot2::element_text(angle=0),
                 legend.position = "none",
                 axis.title=element_text(size=12, face="bold"),
                 plot.title=element_text(size=18),
                 plot.subtitle = element_text(size=14),
                 strip.text = element_text(size=10, face="bold")) +
  ggplot2::scale_x_continuous(limits = c(0, 1), breaks=seq(0, 1, 0.2)) +
  ggplot2::facet_grid(atool~., scales="free_y", space="free") +
  ggplot2::labs(title="Site 3: Indicator and Average Comparisons",
                subtitle="Faceted by assessment tool",
                x="indicator score",
                y="",
                color="assessment tool") +
ggplot2::scale_colour_manual(values=cols)

plot1 + ggplot2::geom_vline(ggplot2::aes(xintercept = score),
                            vline.data1, linetype="solid",
                            color="black", size=2)

#summarize most-disturbed sites via table
final %>%
  mutate(atool = ifelse(indicator=="New LDI", "LDI", as.character(atool)),
         atoolf = factor(atool,
                         levels = c("PFC", "NRCS", "EIA", "DEQ",
                                    "HGM", "MWAM", "LDI")),
         score = ifelse(indicator == "New LDI", 1-(score/8.97), score)) %>%
  filter(grouping=="most disturbed" & (indicator=="average" |
                                       indicator=="New LDI")) %>%
  select(., site, atoolf, score) %>%
  group_by(site, atoolf, score) %>%
  dplyr::rename(`assessment tool`=atoolf) %>%
  pander(caption="Mostly disturbed sites (1, 2, 5, 11, 14, 15 and 16) overall site scores by assessment
final %>%
  filter(grouping=="most disturbed" & (indicator=="average" |
                                       indicator == "New LDI")) %>%
  mutate(atool = ifelse(indicator=="New LDI", "LDI",
                        as.character(atool)),
         atoolf = factor(atool,
                         levels = c("LDI", "MWAM", "HGM",
                                    "DEQ", "EIA","NRCS", "PFC"))) %>%
  select(., atoolf, score) %>%
  group_by(atoolf) %>%
  skimr::skim() %>%
  dplyr::select(atoolf, numeric.mean, numeric.p0, numeric.p100) %>%
  mutate(`assessment tool`=atoolf,
         `mean average score`=round(numeric.mean, 2),
         `min score` = round(numeric.p0, 2),
         `max score`=round(numeric.p100, 2)) %>%
```

```
  dplyr::select(`assessment tool`,
                `mean average score`, `min score`, `max score`) %>%
  mutate(`score range` = round(`max score`- `min score`, 2)) %>%
  pander(caption = "Summary statistics across the most disturbed sites by assessment tool. Average overa
#plot mostly intact sites by score and tool
#create vertical line data with LDI scores for each site
LDIgroup3 <- p2 %>% dplyr::filter(indicator=="New LDI") %>%
  tidyr::gather(key="site", value="score", c(3:20)) %>%
  mutate(score=1-(score/8.97)) %>% dplyr::select(-atool) %>%
  filter(site=="Site 1" | site == "Site 2" | site=="Site 5" | site=="Site 11" |
           site=="Site 14" | site=="Site 15" | site=="Site 16") %>%
  mutate(sitef = factor(site,
                        levels=c("Site 15", "Site 11", "Site 1",
                                 "Site 14", "Site 2", "Site 16","Site 5")))

p <- final %>%
  filter(grouping=="most disturbed" & indicator=="average") %>%
    mutate(atoolf = factor(atool,
                           levels = c("PFC", "NRCS", "EIA",
                                      "DEQ", "HGM", "MWAM", "LDI"))) %>%
  mutate(sitef = factor(site,
                        levels=c("Site 15", "Site 11", "Site 1",
                                 "Site 14", "Site 2", "Site 16","Site 5"))) %>%
  ggplot2::ggplot(ggplot2::aes(y=atoolf, x=score, color=atoolf)) +
  ggplot2::geom_point(ggplot2::aes(y=atoolf, x=score, color=atoolf), size=6) +
  ggplot2::theme(axis.text.y = ggplot2::element_text(angle=0, size=12,
                                                     face="bold", color="black"),
                 axis.text.x = ggplot2::element_text(angle=0, size=12,
                                                     face="bold", color="black"),
                 strip.text.y = element_text(size = 10,
                                             colour = "black", angle = 90),
                 legend.position = "none",
                 axis.title.x=element_text(size=14, face="bold"),
                 plot.title=element_text(size=20)) +
  ggplot2::scale_x_continuous(limits = c(0, 1), breaks=seq(0, 1, 0.1)) +
  ggplot2::facet_grid(sitef~.) +
  ggplot2::labs(title="Most Disturbed Sites (15, 11, 1, 14, 2, 16, 5): Overall Score by Assessment Tool
                x="overall (average) score",
                y="",
                color="assessment tool")
#use this code if LDI are rescaled to the opposite direction and on a scale 0-1
p + ggplot2::geom_vline(ggplot2::aes(xintercept = score),
                        LDIgroup3, linetype="solid",
                        color="black", size=2)

#plot site 5 comparisons

#create vertical line data with averages for each tool
averages <- p_long %>% dplyr::filter(indicator=="average") %>%
  dplyr::select(c("atool", "score", "site"))
#select for specific sites for various graphics
vline.data1 <- averages %>% dplyr::filter(site=="Site 5") %>%
  dplyr::select(c("atool", "score"))
```

```
#create color palette
cols <- RColorBrewer::brewer.pal(n=6, "Dark2")

plot1 <- p_long %>% dplyr::filter(., site == 'Site 5' &
                                    !(indicator=="average") &
                                    !(indicator=="New LDI")) %>%
  ggplot2::ggplot(ggplot2::aes(y=indicator, x=score, color=atool)) +
  ggplot2::geom_point(ggplot2::aes(y=indicator, x=score, color=atool), size=4) +
  ggplot2::theme(axis.text.y = ggplot2::element_text(angle=0),
                 legend.position = "none") +
  ggplot2::theme(axis.text.y = ggplot2::element_text(angle=0),
                 legend.position = "none",
                 axis.title=element_text(size=12, face="bold"),
                 plot.title=element_text(size=18),
                 plot.subtitle = element_text(size=14),
                 strip.text = element_text(size=10, face="bold")) +
  ggplot2::facet_grid(atool~., scales="free_y", space="free") +
  ggplot2::labs(title="Site 5: Indicator and Average Comparisons",
                subtitle="Faceted by assessment tool",
                x="indicator score",
                y="",
                color="assessment tool") +
ggplot2::scale_colour_manual(values=cols)

plot1 + ggplot2::geom_vline(ggplot2::aes(xintercept = score),
                            vline.data1, linetype="solid",
                            color="black", size=2)

#plot all sites within given grouping faceted by grouping
#create vertical line data with LDI scores for each site
final %>%  filter(indicator=="New LDI" | indicator=="average") %>%
  mutate(LDIscore = ave(score, site, FUN = max),
         LDIscore_stand = 1-(LDIscore/8.97),
         atool = ifelse(indicator=="New LDI", "LDI", as.character(atool)),
         atoolf = factor(atool,
                         levels = c("PFC", "NRCS",
                                    "EIA", "DEQ", "HGM", "MWAM", "LDI")),
         score = ifelse(indicator == "New LDI", 1-(score/8.97), score),
         groupingf = factor(grouping,
                            levels=c("mainly intact", "mid-range","most disturbed")),
         sitef = as.numeric(str_remove(site, "Site "))) %>%
  #change order of groupings to be in logical order
  #change factor order of atool
  ggplot2::ggplot(ggplot2::aes(y=atoolf, x=score)) +
  ggplot2::geom_point(aes(y=atoolf, x=score, color=LDIscore_stand),
                      position=position_dodge(width=0.4), size=6) +
  geom_text_repel(mapping=aes(x=score, y=atoolf,label=sitef),
    size=5, box.padding = unit(0.5, "lines")) +
  theme_bw() +
  ggplot2::theme(axis.text.y = ggplot2::element_text(angle=0, size=12,
                                                     face="bold", color="black"),
                 axis.text.x = ggplot2::element_text(angle=0, size=12,
                                                     face="bold", color="black"),
```

```
                strip.text.y = element_text(size = 12,
                                            colour = "black", angle = 90),
                axis.title.x=element_text(size=14, face="bold"),
                plot.title=element_text(size=20),
                legend.title = element_text(size=13, face="bold"),
                legend.text = element_text(size=12),
                legend.position = c(0.1, 0.17),
                legend.background = element_rect(fill="lightgrey",
                                  size=1, linetype="solid",
                                  colour ="black"),
                legend.key.size =  unit(1, "cm"),
                legend.box="horizontal",
                axis.ticks = element_blank(),
                panel.grid.major = element_line(colour="grey70", size=0.3)) +
  ggplot2::scale_x_continuous(limits = c(0, 1), breaks=seq(0, 1, 0.1)) +
  scale_color_viridis(option="B", begin = 0, end = 1, direction = -1,
                      limits=c(0,1), breaks=c(0, 0.5, 1)) +
  ggplot2::facet_grid(groupingf~.) +
  ggplot2::labs(title="Overall Site Scores by Ecological Health: Most Disturbed, Mid-range and Mainly In
                x="overall (average) score",
                y="",
                color="LDI site score")

#set themes for all plots
fills <- viridis::scale_fill_viridis(option="B", direction = -1, limits=c(0,1))

themes <- ggplot2::theme(axis.text.y = element_text(angle=0,
                                                    size=12,
                                                    face="bold", color="black"),
                axis.text.x = element_text(angle=0,
                                           size=12, face="bold", color="black"),
                axis.title.x=element_blank(),
                axis.title.y=element_blank(),
                plot.title=element_text(size=20, face="bold", hjust = 0.5),
                legend.position="none")

orderDEQ <-
  p2 %>%tidyr::gather(key="site", value="score", c(3:20)) %>%
  mutate(atool = ifelse(indicator=="New LDI", "New LDI",
                        as.character(atool))) %>%
  filter(atool=="DEQ") %>%
  mutate(site = str_remove(site, "Site")) %>%
  group_by(site) %>% filter(indicator=="average") %>%
  arrange(desc((as.numeric(as.character(score)))))
orderDEQ$rank <- 1:nrow(orderDEQ)

tiledeq <- p2 %>% tidyr::gather(key="site", value="score", c(3:20)) %>%
  mutate(atool = ifelse(indicator=="New LDI", "New LDI", as.character(atool)),
         score=as.numeric(as.character(score)),
         site = str_remove(site, "Site")) %>%
  filter(atool=="DEQ") %>%
  mutate(indicatorF = fct_relevel(indicator, "average",
                                  "landscape context", "veg structure",
```

```
                                    "physiochemical", "hydrology")) %>%
  left_join(., (orderDEQ %>% select(., c(rank, site))), by="site", keep=FALSE) %>%
  ggplot() +
  geom_tile(aes(x=fct_reorder(as.factor(site), rank),
                y=indicatorF,
                fill=score)) +
  fills +
  themes +
  labs(title="DEQ",
       x="",
       y="")

orderMTNHP <-
  p2 %>% tidyr::gather(key="site", value="score", c(3:20)) %>%
  mutate(atool = ifelse(indicator=="New LDI", "New LDI", as.character(atool))) %>%
  filter(atool=="EIA") %>%
  mutate(site = str_remove(site, "Site")) %>%
  group_by(site) %>% filter(indicator=="average") %>%
  arrange(desc((as.numeric(as.character(score)))))
orderMTNHP$rank <- 1:nrow(orderMTNHP)

tilemtnhp <- p2 %>% tidyr::gather(key="site", value="score", c(3:20)) %>%
  mutate(atool = ifelse(indicator=="New LDI", "New LDI", as.character(atool)),
         score=as.numeric(as.character(score)),
         site = str_remove(site, "Site")) %>%
  filter(atool=="EIA") %>%
  mutate(indicatorF = fct_relevel(indicator, "average", "landscape",
                                  "vegetation", "water quality", "hydrology")) %>%
  left_join(., (orderMTNHP %>% select(., c(rank, site))),
            by="site", keep=FALSE) %>%
  ggplot() +
  geom_tile(aes(x=fct_reorder(as.factor(site), rank),
                y=indicatorF,
                fill=score)) +
  fills +
  themes +
  labs(title="EIA",
       x="",
       y="")

orderHGM <-
  p2 %>% tidyr::gather(key="site", value="score", c(3:20)) %>%
  mutate(atool = ifelse(indicator=="New LDI", "New LDI", as.character(atool))) %>%
  filter(atool=="HGM") %>%
  mutate(site = str_remove(site, "Site")) %>%
  group_by(site) %>% filter(indicator=="average") %>%
  arrange(desc((as.numeric(as.character(score)))))
orderHGM$rank <- 1:nrow(orderHGM)

tilehgm <- p2 %>% tidyr::gather(key="site", value="score", c(3:20)) %>%
  mutate(atool = ifelse(indicator=="New LDI", "New LDI", as.character(atool)),
         score=as.numeric(as.character(score)),
         site = str_remove(site, "Site")) %>%
```

```
  filter(atool=="HGM") %>%
  left_join(., (orderHGM %>% select(., c(rank, site))), by="site", keep=FALSE) %>%
  mutate(indicatorF = fct_relevel(indicator,  "average",
                                  "vertebrate habitat", "aquatic invert food web",
                                  "plant community", "in/organic particle retention",
                                  "organic C gen/export", "nutrient cycling",
                                  "surface-groundwater",
                                  "floodplain int/connectivity")) %>%
  ggplot() +
  geom_tile(aes(x=fct_reorder(as.factor(site), rank),
                y=indicatorF,
                fill=score)) +
  fills +
  themes +
  labs(title="HGM",
       x="",
       y="",
       fill="")

nrcs <- readxl::read_excel("assessment_033121.xlsx", sheet = "Mushed values") %>%
  dplyr::slice(19:23) %>%
  dplyr::rename(indicator = Question) %>%
  mutate(indicator = ifelse(indicator=="Geomorphic Considerations", "geomorphic",
                            ifelse(indicator=="Vegetation Considerations",
                                   "vegetation",
                            ifelse(indicator=="wood_browse", "utilization",
                            ifelse(indicator=="fldpln_erg", "floodplain",
                            ifelse(indicator=="Rating", "average",
                                   indicator))))),
         indicatorF = fct_relevel(indicator, "average", "utilization",
                                  "vegetation", "geomorphic", "floodplain"))

orderNRCS <-
  nrcs %>% tidyr::gather(key="site", value="score", c(2:19)) %>%
  mutate(site = str_remove(site, "Site")) %>%
  group_by(site) %>% filter(indicatorF=="average") %>%
  arrange(desc((as.numeric(as.character(score)))))
orderNRCS$rank <- 1:nrow(orderNRCS)

tilenrcs <- nrcs %>% tidyr::gather(key="site", value="score", c(2:19)) %>%
  mutate(score=as.numeric(as.character(score)), site = str_remove(site, "Site")) %>%
  left_join(., (orderNRCS %>% select(., c(rank, site))), by="site", keep=FALSE) %>%
  ggplot() +
  geom_tile(aes(x=fct_reorder(as.factor(site), rank),
                y=indicatorF,
                fill=score)) +
  fills +
  themes +
  labs(title="NRCS",
       x="",
       y="",
       fill="")
```

```
mdot <- readxl::read_excel("assessment_033121.xlsx", sheet = "Mushed values") %>%
  dplyr::slice(2:9) %>%
  dplyr::rename(indicator = Question) %>%
  mutate(indicator = ifelse(indicator=="hab_state_rate", "habitat state S1/S2/S3",
                            ifelse(indicator=="General habitat", "gen habitat",
                            ifelse(indicator=="flood_attn", "flood attenuation",
                            ifelse(indicator=="Water Storage Exchange",
                                   "water storage/exchange",
                            ifelse(indicator=="Sediment source and storage",
                                   "sed source/storage",
                            ifelse(indicator=="food_chain_suport",
                                   "food chain support",
                            ifelse(indicator=="Total Score", "average",
                            ifelse(indicator=="Uniqueness", "uniqueness",
                                   indicator)))))))),
         indicatorF = fct_relevel(indicator, "average", "uniqueness",
                                  "gen habitat", "habitat state S1/S2/S3",
                                  "food chain support", "sed source/storage",
                                  "water storage/exchange", "flood attenuation"))

orderMDOT <-
  mdot %>% tidyr::gather(key="site", value="score", c(2:19)) %>%
  mutate(site = str_remove(site, "Site")) %>%
  group_by(site) %>% filter(indicatorF=="average") %>%
  arrange(desc((as.numeric(as.character(score)))))
orderMDOT$rank <- 1:nrow(orderMDOT)

tilemdot <- mdot %>% tidyr::gather(key="site", value="score", c(2:19)) %>%
  mutate(score=as.numeric(as.character(score)),
         site = str_remove(site, "Site")) %>%
  left_join(., (orderMDOT %>% select(., c(rank, site))),
            by="site", keep=FALSE) %>%
  ggplot() +
  geom_tile(aes(x=fct_reorder(as.factor(site), rank),
                y=indicatorF,
                fill=score)) +
  fills +
  themes +
  labs(title="MWAM",
       x="",
       y="",
       fill="")

blm <- readxl::read_excel("assessment_033121.xlsx", sheet = "BLM-REvised") %>%
  dplyr::slice(76:80) %>%
  dplyr::rename(indicator = Question) %>%
  mutate(indicator = ifelse(indicator=="Hydrology", "hydrology",
                            ifelse(indicator=="Vegetation", "vegetation",
                            ifelse(indicator=="Geomorphology", "geomorphology",
                            ifelse(indicator=="Carbon Export", "carbon",
                            ifelse(indicator=="Rating(just average)", "average",
                                   indicator))))),
         indicatorF = fct_relevel(indicator, "average", "vegetation",
```

```
                                  "carbon", "geomorphology", "hydrology"))

orderBLM <-
  blm %>% tidyr::gather(key="site", value="score", c(2:19)) %>%
  mutate(site = str_remove(site, "Site")) %>%
  group_by(site) %>% filter(indicatorF=="average") %>%
  arrange(desc((as.numeric(as.character(score)))))
orderBLM$rank <- 1:nrow(orderBLM)

tileblm <- blm %>% tidyr::gather(key="site", value="score", c(2:19)) %>%
  mutate(score=as.numeric(as.character(score)),
         site = str_remove(site, "Site")) %>%
  left_join(., (orderBLM %>% select(., c(rank, site))),
            by="site", keep=FALSE) %>%
  ggplot() +
  geom_tile(aes(x=fct_reorder(as.factor(site), rank),
                y=indicatorF,
                fill=score)) +
  fills +
  themes +
  labs(title="PFC",
       x="",
       y="",
       fill="")

legendcor <- blm %>% tidyr::gather(key="site", value="score", c(2:19)) %>%
  mutate(score=as.numeric(as.character(score))) %>%
  left_join(., (orderBLM %>% select(., c(rank, site))),
            by="site", keep=FALSE) %>%
  ggplot() +
  geom_tile(aes(x=fct_reorder(as.factor(site), rank),
                y=indicatorF,
                fill=score)) +
  scale_fill_continuous(guide =guide_legend(title.position="top")) +
  scale_fill_viridis_c(option="B", begin = 0, end = 1, direction = -1,
                       limits=c(0,1)) +
  ggplot2::theme(legend.position = "bottom",
                 legend.box="horizontal",
                 legend.text = element_text(colour="black", size = 12,
                                            face="bold"),
                 legend.title = element_text(color="black", size = 12,
                                             face="bold", hjust = 0.5),
                 legend.key.size = unit(0.5, "cm"),
                 legend.key.width = unit(3.5, "cm")) +
  labs(title="",
       x="",
       y="",
       fill="")

#extract legend from ggplot function
extract_legend <- function(my_ggp) {
  step1 <- ggplot_gtable(ggplot_build(my_ggp))
  step2 <- which(sapply(step1$grobs, function(x) x$name) == "guide-box")
```

```
  step3 <- step1$grobs[[step2]]
  return(step3)
}
#run function to extract
shared_legend <- extract_legend(legendcor)

grid.arrange(arrangeGrob(tilemdot, tilehgm, tiledeq,
                         tilemtnhp, tilenrcs, tileblm,
                         ncol=2), shared_legend,
             nrow=3, heights=c(10, 1, 10))
themes <- ggplot2::theme(axis.text.y = element_text(size=12,
                                                    face="bold", color="black"),
                 axis.text.x = element_text(size=12, face="bold", color="black"),
                 axis.title.x=element_blank(),
                 axis.title.y=element_blank(),
                 plot.title=element_text(size=20, face="bold", hjust = 0.5),
                 legend.position="none")

MDOTcor <- readxl::read_excel("assessment_033121.xlsx", sheet = "Mushed values") %>%
  dplyr::slice(2:9) %>%
  dplyr::rename(indicator = Question) %>%
  mutate(indicator = ifelse(indicator=="hab_state_rate", "habitat state S1/S2/S3",
                            ifelse(indicator=="General habitat", "gen habitat",
                            ifelse(indicator=="flood_attn", "flood attenuation",
                            ifelse(indicator=="Water Storage Exchange",
                                   "water storage/exchange",
                            ifelse(indicator=="Sediment source and storage",
                                   "sed source/storage",
                            ifelse(indicator=="food_chain_suport",
                                   "food chain support",
                            ifelse(indicator=="Total Score", "average",
                            ifelse(indicator=="Uniqueness", "uniqueness",
                                   indicator)))))))))) %>%
  tidyr::gather(key="site", value="score", c(2:19)) %>%
  pivot_wider(., names_from=indicator, values_from=score) %>%
  select(-c(average, site)) %>%
  cor(.) %>%
  ggcorrplot::ggcorrplot(., method="circle", type="upper",
                         show.legend = FALSE, legend.title="",
                         title = "MWAM", colors=c("red", "white", "blue")) +
  themes

HGMcor <- p2 %>%
  tidyr::gather(key="site", value="score", c(3:20)) %>%
  mutate(atool = ifelse(indicator=="New LDI", "New LDI", as.character(atool)),
         score=as.numeric(as.character(score))) %>%
  filter(atool=="HGM") %>%
  pivot_wider(., names_from=indicator, values_from=score) %>%
  select(-c(atool, average, site)) %>%
  cor(.) %>%
  ggcorrplot::ggcorrplot(., method="circle", type="upper",
                         show.legend = FALSE, legend.title="",
                         title = "HGM Approach",
```

```
                          colors=c("red", "white", "blue")) +
  themes

DEQcor <- p2 %>%
  tidyr::gather(key="site", value="score", c(3:20)) %>%
  mutate(atool = ifelse(indicator=="New LDI", "New LDI",
                        as.character(atool)),
         score=as.numeric(as.character(score))) %>%
  filter(atool=="DEQ") %>%
  pivot_wider(., names_from=indicator, values_from=score) %>%
  select(-c(atool, average, site)) %>%
  cor(.) %>%
  ggcorrplot::ggcorrplot(., method="circle", type="upper",
                         show.legend = FALSE, legend.title="",
                         title = "DEQ", colors=c("red", "white", "blue")) +
  themes

MTNHPcor <- p2 %>%
  tidyr::gather(key="site", value="score", c(3:20)) %>%
  mutate(atool = ifelse(indicator=="New LDI", "New LDI", as.character(atool)),
         score=as.numeric(as.character(score))) %>%
  filter(atool=="EIA") %>%
  pivot_wider(., names_from=indicator, values_from=score) %>%
  select(-c(atool, average, site)) %>%
  cor(.) %>%
  ggcorrplot::ggcorrplot(., method="circle", type="upper",
                         show.legend = FALSE, legend.title="",
                         title = "EIA", colors=c("red", "white", "blue")) +
  themes

NRCScor <- readxl::read_excel("assessment_033121.xlsx", sheet = "Mushed values") %>%
  dplyr::slice(19:23) %>%
  dplyr::rename(indicator = Question) %>%
  mutate(indicator = ifelse(indicator=="Geomorphic Considerations",
                            "geomorphic",
                            ifelse(indicator=="Vegetation Considerations",
                                   "vegetation",
                            ifelse(indicator=="wood_browse", "utilization",
                            ifelse(indicator=="fldpln_erg", "floodplain",
                            ifelse(indicator=="Rating", "average",
                                   indicator)))))) %>%
  tidyr::gather(key="site", value="score", c(2:19)) %>%
  pivot_wider(., names_from=indicator, values_from=score) %>%
  select(-c(average, site)) %>%
  cor(.) %>%
  ggcorrplot::ggcorrplot(., method="circle", type="upper",
                         show.legend = FALSE, legend.title="",
                         title = "NRCS", colors=c("red", "white", "blue")) +
  themes

BLMcor <- readxl::read_excel("assessment_033121.xlsx", sheet = "BLM-REvised") %>%
  dplyr::slice(76:80) %>%
  dplyr::rename(indicator = Question) %>%
```

```
  mutate(indicator = ifelse(indicator=="Hydrology", "hydrology",
                            ifelse(indicator=="Vegetation", "vegetation",
                            ifelse(indicator=="Geomorphology", "geomorphology",
                            ifelse(indicator=="Carbon Export", "carbon",
                            ifelse(indicator=="Rating(just average)",
                                   "average", indicator)))))) %>%
  tidyr::gather(key="site", value="score", c(2:19)) %>%
  pivot_wider(., names_from=indicator, values_from=score) %>%
  select(-c(average, site)) %>%
  mutate_all(., as.numeric) %>%
  cor(.) %>%
  ggcorrplot::ggcorrplot(., method="circle", type="upper",
                         show.legend = FALSE, legend.title="",
                         title = "PFC", colors=c("red", "white", "blue")) +
  themes

legendcor <- p2 %>%
  tidyr::gather(key="site", value="score", c(3:20)) %>%
  mutate(atool = ifelse(indicator=="New LDI", "New LDI", as.character(atool)),
         score=as.numeric(as.character(score))) %>%
  filter(atool=="DEQ") %>%
  pivot_wider(., names_from=indicator, values_from=score) %>%
  select(-c(atool, average, site)) %>%
  cor(.) %>%
  ggcorrplot::ggcorrplot(., method="circle", type="upper",
                         legend.title="", title = "BLM",
                         colors=c("red", "white", "blue")) +
  theme(legend.position = "bottom",
        legend.box="horizontal",
        legend.text = element_text(colour="black", size = 14, face="bold"),
        legend.key.size = unit(1, "cm"),
        legend.key.width = unit(2, "cm"))
#extract legend from ggplot function
extract_legend <- function(my_ggp) {
  step1 <- ggplot_gtable(ggplot_build(my_ggp))
  step2 <- which(sapply(step1$grobs, function(x) x$name) == "guide-box")
  step3 <- step1$grobs[[step2]]
  return(step3)
}
#run function to extract
shared_legend <- extract_legend(legendcor)

grid.arrange(arrangeGrob(MDOTcor, HGMcor, DEQcor,
                         MTNHPcor, NRCScor, BLMcor, ncol=2),
             shared_legend, nrow=3, heights=c(10, 1, 2))

themes <- ggplot2::theme(axis.text.y = element_text(angle=0, size=12,
                                                    face="bold", color = "black"),
                         axis.title.y = element_blank(),
                 axis.text.x = element_text(angle=0, size=12, face="bold"),
                 axis.title.x=element_blank(),
                 plot.title=element_text(size=20, face="bold", hjust = 0.5),
                 legend.position="none")
```

```
scalesx <- scale_x_continuous(seq(0,1, by=0.25), limits=c(0,1.01))
scalesfill <- scale_fill_gradient2(name = "Loading",
                        high= "#0C6291", mid = "#FBFEF9", low = "#A63446",
                        midpoint=0, guide=F)
nrcs.factor <- readxl::read_excel("assessment_033121.xlsx",
                                  sheet = "Mushed values") %>%
  dplyr::slice(19:23) %>%
  dplyr::rename(indicator = Question) %>%
  mutate(indicator = ifelse(indicator=="Geomorphic Considerations",
                            "geomorphic",
                            ifelse(indicator=="Vegetation Considerations",
                                   "vegetation",
                            ifelse(indicator=="wood_browse", "utilization",
                            ifelse(indicator=="fldpln_erg", "floodplain",
                            ifelse(indicator=="Rating", "average",
                                   indicator)))))) %>%
  tidyr::gather(key="site", value="score", c(2:19)) %>%
  pivot_wider(., names_from=indicator, values_from=score) %>%
  select(-c(average, site)) %>%
  cor(.) %>%
  psych::fa(r=., nfactors=1, rotate="none", fm="minres")
nrcs.df <- data.frame(index = factor(c("geomorphic", "vegetation",
                                       "utilization", "floodplain"),
                                     levels=c("utilization", "vegetation",
                                              "geomorphic", "floodplain")),
                      loadings = c(as.vector(nrcs.factor$loadings)))
n <- ggplot(nrcs.df, aes(y=index, x=abs(loadings), fill=loadings)) +
  geom_bar(stat="identity") +
  geom_text(aes(label = round(loadings, 2), hjust = 1.5), size = 7, face="bold") +
  scalesfill +
  themes +
  scalesx +
  labs(title = "NRCS")

mdot.factor <- readxl::read_excel("assessment_033121.xlsx", sheet = "Mushed values") %>%
  dplyr::slice(2:9) %>%
  dplyr::rename(indicator = Question) %>%
  mutate(indicator = ifelse(indicator=="hab_state_rate", "habitat state S1/S2/S3",
                            ifelse(indicator=="General habitat", "gen habitat",
                            ifelse(indicator=="flood_attn", "flood attenuation",
                            ifelse(indicator=="Water Storage Exchange",
                                   "water storage/exchange",
                            ifelse(indicator=="Sediment source and storage",
                                   "sed source/storage",
                            ifelse(indicator=="food_chain_suport",
                                   "food chain support",
                            ifelse(indicator=="Total Score", "average",
                            ifelse(indicator=="Uniqueness", "uniqueness",
                                   indicator)))))))))) %>%
  tidyr::gather(key="site", value="score", c(2:19)) %>%
  pivot_wider(., names_from=indicator, values_from=score) %>%
  select(-c(average, site)) %>%
  cor(.) %>%
```

```
  psych::fa(r=., nfactors=1, rotate="none", fm="minres")
mdot.df <- data.frame(index = factor(c("habitat state S1/S2/S3",
                                       "gen habitat", "water storage/exchange",
                                       "sed source/storage", "flood attenuation",
                                       "food chain support", "uniqueness"),
                                     levels=c("uniqueness", "gen habitat",
                                              "habitat state S1/S2/S3",
                                              "food chain support",
                                              "sed source/storage",
                                              "water storage/exchange",
                                              "flood attenuation")),
                      loadings = c(as.vector(mdot.factor$loadings)))
md <- ggplot(mdot.df, aes(y=index, x=abs(loadings), fill=loadings)) +
  geom_bar(stat="identity") +
  geom_text(aes(label = round(loadings, 2), hjust = 1.5), size = 7, face="bold") +
  scalesfill +
  themes +
  scalesx +
  labs(title = "MWAM")

blm.factor <- readxl::read_excel("assessment_033121.xlsx",
                                 sheet = "BLM-REvised") %>%
  dplyr::slice(76:80) %>%
  dplyr::rename(indicator = Question) %>%
  mutate(indicator = ifelse(indicator=="Hydrology", "hydrology",
                            ifelse(indicator=="Vegetation", "vegetation",
                            ifelse(indicator=="Geomorphology", "geomorphology",
                            ifelse(indicator=="Carbon Export", "carbon",
                            ifelse(indicator=="Rating(just average)",
                                   "average", indicator))))) %>%
  tidyr::gather(key="site", value="score", c(2:19)) %>%
  pivot_wider(., names_from=indicator, values_from=score) %>%
  select(-c(average, site)) %>%
  mutate_all(., as.numeric) %>%
  cor(.) %>%
  psych::fa(r=., nfactors=1, rotate="none", fm="minres")
blm.df <- data.frame(index = factor(c("hydrology", "vegetation",
                                      "geomorphology", "carbon"),
                                    levels= c("vegetation", "carbon",
                                              "geomorphology", "hydrology")),
                     loadings = c(as.vector(blm.factor$loadings)))
b <- ggplot(blm.df, aes(y=index, x=abs(loadings), fill=loadings)) +
  geom_bar(stat="identity") +
  geom_text(aes(label = round(loadings, 2), hjust = 1.5), size = 7, face="bold") +
  scalesfill +
  themes +
  scalesx +
  labs(title = "PFC")

deq.factor <- p2 %>%
  tidyr::gather(key="site", value="score", c(3:20)) %>%
  mutate(atool = ifelse(indicator=="New LDI", "New LDI", as.character(atool)),
         score=as.numeric(as.character(score))) %>%
```

```
  filter(atool=="DEQ") %>%
  pivot_wider(., names_from=indicator, values_from=score) %>%
  select(-c(atool, average, site)) %>%
  cor(.) %>%
  psych::fa(r=., nfactors=1, rotate="none", fm="minres")
deq.df <- data.frame(index = factor(c("landscape context",
                                      "veg structure", "physiochemical",
                                      "hydrology"),
                                    levels = c("landscape context", "veg structure",
                                               "physiochemical", "hydrology")),
                     loadings = c(as.vector(deq.factor$loadings)))
d <- ggplot(deq.df, aes(y=as.factor(index), x=abs(loadings), fill=loadings)) +
  geom_bar(stat="identity") +
  geom_text(aes(label = round(loadings, 2), hjust = 1.5), size = 7, face="bold") +
  scalesfill +
  themes +
  scalesx +
  labs(title = "DEQ")

hgm.factor <- p2 %>%
  tidyr::gather(key="site", value="score", c(3:20)) %>%
  mutate(atool = ifelse(indicator=="New LDI", "New LDI", as.character(atool)),
         score=as.numeric(as.character(score))) %>%
  filter(atool=="HGM") %>%
  pivot_wider(., names_from=indicator, values_from=score) %>%
  select(-c(atool, average, site)) %>%
  cor(.) %>%
  psych::fa(r=., nfactors=1, rotate="none", fm="minres")
  hgm.df <- data.frame(index = factor(c("surface-groundwater",
                                        "nutrient cycling",
                                        "in/organic particle retention",
                                        "organic C gen/export", "plant community",
                                        "aquatic invert food web", "vertebrate habitat",
                                        "floodplain int/connectivity"),
                                      levels=c("vertebrate habitat",
                                               "aquatic invert food web",
                                               "plant community",
                                               "in/organic particle retention",
                                               "organic C gen/export",
                                               "nutrient cycling","surface-groundwater",
                                               "floodplain int/connectivity")),
                       loadings = c(as.vector(hgm.factor$loadings)))
h <- ggplot(hgm.df, aes(y=index, x=abs(loadings), fill=loadings)) +
  geom_bar(stat="identity") +
  geom_text(aes(label = round(loadings, 2), hjust = 1.5), size = 7, face="bold") +
  scalesfill +
  themes +
  scalesx +
  labs(title = "HGM")

mtnhp.factor <- p2 %>%
  tidyr::gather(key="site", value="score", c(3:20)) %>%
  mutate(atool = ifelse(indicator=="New LDI", "New LDI", as.character(atool)),
```

```
        score=as.numeric(as.character(score))) %>%
  filter(atool=="EIA") %>%
  pivot_wider(., names_from=indicator, values_from=score) %>%
  select(-c(atool, average, site)) %>%
  cor(.) %>%
  psych::fa(r=., nfactors=1, rotate="none", fm="minres")
  mtnhp.df <- data.frame(index = factor(c("landscape", "vegetation",
                                          "water quality", "hydrology"),
                                        levels=c("landscape", "vegetation",
                                                 "water quality", "hydrology")),
                         loadings = c(as.vector(mtnhp.factor$loadings)))
m <- ggplot(mtnhp.df, aes(y=index, x=abs(loadings), fill=loadings)) +
  geom_bar(stat="identity") +
  geom_text(aes(label = round(loadings, 2), hjust = 1.5), size = 7, face="bold") +
  scalesfill +
  themes +
  scalesx +
  labs(title = "EIA")
legendcor <- ggplot(mtnhp.df, aes(y=index, x=abs(loadings), fill=loadings)) +
  geom_bar(stat="identity") +
  scale_x_continuous(seq(0,1, by=0.25), limits=c(0,1)) +
  scale_fill_gradient2(name = "",
                       high= "#0C6291", mid = "#FBFEF9", low = "#A63446",
                       midpoint=0,
                       breaks = c(-1, -0.5, 0, 0.5, 1),
                       labels = c("-1", "-0.5", "0", "0.5", "1"),
                       limits = c(-1, 1)) +
  theme(legend.position = "bottom",
        legend.box="horizontal",
        legend.text = element_text(colour="black", size = 14, face="bold"),
        legend.key.size = unit(1, "cm"),
        legend.key.width = unit(2, "cm"),
        legend.title.align = 0.5)

#extract legend from ggplot function
extract_legend <- function(my_ggp) {
  step1 <- ggplot_gtable(ggplot_build(my_ggp))
  step2 <- which(sapply(step1$grobs, function(x) x$name) == "guide-box")
  step3 <- step1$grobs[[step2]]
  return(step3)
}
#run function to extract
shared_legend <- extract_legend(legendcor)

grid.arrange(arrangeGrob(md, h, d, m, n, b, ncol=2),
             shared_legend, nrow=3, heights=c(10, 1, 2))
options(width = 60)
```