



# Choosing the Right Tool: A Comparative Study of Wetland Assessment Approaches

William J. Kleindl<sup>1</sup>  · Sarah P. Church<sup>2</sup> · Mark C. Rains<sup>3</sup> · Rachel Ulrich<sup>4</sup>

Received: 17 May 2023 / Accepted: 16 March 2024  
© The Author(s) 2024

## Abstract

There are over 700 aquatic ecological assessment approaches across the globe that meet specific institutional goals. However, in many cases, multiple assessment tools are designed to meet the same management need, resulting in a confusing array of overlapping options. Here, we look at six riverine wetland assessments currently in use in Montana, USA, and ask which tool (1) best captures the condition across a disturbance gradient and (2) has the most utility to meet the regulatory or management needs. We used descriptive statistics to compare wetland assessments ( $n = 18$ ) across a disturbance gradient determined by a landscape development intensity. Factor analysis showed that many of the tools had internal metrics that did not correspond well with overall results, hindering the tool's ability to act as designed. We surveyed regional wetland managers ( $n = 56$ ) to determine the extent of their use of each of the six tools and how well they trusted the information the assessment tool provided. We found that the Montana Wetland Assessment Methodology best measured the range of disturbance and had the highest utility to meet Clean Water Act (CWA § 404) needs. Montana Department of Environmental Quality was best for the CWA § 303(d) & 305(b) needs. The US Natural Resources Conservation Service's Riparian Assessment Tool was the third most used by managers but was the tool that had the least ability to distinguish across a disturbance, followed by the US Bureau of Land Management's Proper Functioning Condition.

**Keywords** Wetland · Riparian · Assessment · Regulatory · Factor analysis

## Introduction

Following a thread of civil law as far back as 500 A.D., running waters have been considered common to all and property of none, and the public benefit of this shared resource within the United States (U.S.) requires that it be preserved and protected (SCOTUS 1892; SCOTUS 1950). These ideas form the basis of federal and state water laws, rules,

and guidance. Central to these is a need to understand the ecological condition of those waters, the potential harm to the public resource that may occur from land-use decisions and other anthropogenic disturbances, and how to mitigate that harm. Ecological monitoring and assessment of these systems' chemical, physical, and biotic components facilitate that central need. However, because of the shifting regulatory requirements at federal and state levels, specific agency monitoring mandates, as well as the changes in assessment approaches over the last several decades, there has been a proliferation of aquatic monitoring and assessment tools leading to a confusing array of potential options (Kusler 2006; Stein et al. 2009). Moreover, novel regulatory tools are emerging that assess elements of aquatic ecosystems that maintain human well-being (aquatic ecosystem services) (e.g., Stelk and Christie 2014; USEPA 2020) that may further add to the proliferation.

Ecological monitoring and assessment are nearly as old as the science of ecology (Hynes 1974), beginning in the early 1900s to estimate water pollution levels (Kolkwitz and Marsson 1908; Bick 1963). Across the 20th century,

---

✉ William J. Kleindl  
william.kleindl@montana.edu

<sup>1</sup> Dept of Land Resources and Environmental Sciences, Montana State University, 334 Leon Johnson Hall, 59717 Bozeman, MT, USA

<sup>2</sup> Dept Earth Sciences, Montana State University, 226 Traphagen Hall, 59717 Bozeman, MT, USA

<sup>3</sup> School of Geosciences, University of South Florida, 33620 Tampa, FL, USA

<sup>4</sup> Statistical Sciences Los Alamos National Lab, 87545 Los Alamos, NM, USA

assessment techniques and tools applied theoretical and empirical advances in ecological sciences to measure the impacts of emergent environmental crises (Cairns Jr and Pratt 1993; Verdonshot 2000). Initiatives to protect the ecological condition of U.S. federal and state waters began with 1970s U.S. environmental regulation and guidance. Thus, began broad iterative advances in aquatic ecological knowledge and regulatory oversight. For example, since the U.S. ‘no net loss’ policy for the management of wetlands of the late 1980s, there has been an increasing trend in the scientific literature focusing on wetland function, condition, and value that iterate with the development of a variety of wetland assessment approaches to measure those elements. Each assessment approach has a general common objective: to evaluate the complex ecological condition of aquatic resources using a finite set of observable biophysical fields and spatial indicators (Stein et al. 2009). The condition is generally determined by how much a particular site diverges from an ideal state across a disturbance gradient (Rappport et al. 1985; Fennessy et al. 2007; Stein et al. 2009). This expression of relative condition informs decision-makers of their management or regulatory needs (Stein et al. 2009).

Despite this shared objective, there is no unified assessment approach, as each is designed to meet specific institutional goals. As a result, there are over 700 qualitative and quantitative aquatic ecological assessment approaches across the globe that address the impacts of direct human activities on systems (e.g., Fennessy et al. 2004; Goodrich et al. 2005; Birk et al. 2012; Wellemeyer et al. 2018; Poikane et al. 2020). This set of assessment tools is comprised of many partially overlapping subsets. Currently, in the U.S., these assessment approaches fall into three broad categories: (1) those that support regulatory obligations of the Clean Water Act (CWA§ 404) to assist in the permitted fill of protected waters; (2) CWA§ 303(d) & 305(b) monitoring, assessing, and reporting of the condition of our nation’s waters; and (3) those that meet resource management mandates of other federal and state agencies. Although management needs should drive the selection of the appropriate assessment approach (Stein et al. 2009), there are cases where there are multiple assessment tools designed to meet the same management need.

Because each organization developed tools to meet its needs, there may be no single ‘best’ tool. Yet, an evaluator may wish to know the most appropriate tool for a specific need. How do we choose the most appropriate assessment approach to meet a management or regulatory need in areas with overlapping wetland function/condition tools? There are only a few examples in the literature for cross-comparing assessment tools (e.g., Gaucherand et al. 2015; Bezombes et al. 2017), and those did not provide a clear path of comparison. We seek to provide a path of comparison

by determining which overlapping biophysical assessment tools (1) best capture conditions across a disturbance gradient and (2) have the most utility to meet the most extensive regulatory need. We do this here, using Montana (U.S.) as a case study and creating a duplicable process for other regions. This analysis intends to serve as a general guideline for assessing the ability of the various tools to describe a set of sample sites.

## Methods

### Assessment Tool Selection

For this case study, we chose floodplain wetlands in the headwaters of the Missouri and Yellowstone Rivers in Montana (U.S.). Within Montana and the four surrounding U.S. intermountain western states, twenty-five ecosystem assessment tools have been developed to measure aquatic health and, in a few cases, some limited measures of ecological service. There are currently twelve wetland assessment tools in Montana, several of which overlap with the neighboring intermountain western states. Of these twelve, we focused only on rapid assessment tools (viz. Fennessy et al. 2007; Kleindl et al. 2010) used in regulatory or resource management of Montana Rocky Mountain riparian wetlands. With these criteria, we chose six Montana wetland assessment approaches to examine which tool best captures a range of conditions across a disturbance gradient and has the most utility to meet the most extensive regulatory need (Table 1).

To address our research questions, we broke the study into two parts: (1) address the tool’s abilities to capture a nuance of conditions across a disturbance gradient and (2) measure the tool’s utility among end-users.

### Range of Conditions

#### Landscape Disturbance Index and Case Study Location

We established a landscape development intensity (LDI) to help select study sites representing a gradient of anthropogenic disturbance. We followed LDI protocols used elsewhere in the U.S. (Brown and Vivas 2005; Wardrop et al. 2016). The LDI establishes a coarse rank of potential wetland study sites according to their underlying anthropogenic land use and establishes the final wetland study sites through an iterative process. We selected our initial sites from National Wetland Inventory (NWI) data (USFWS 2020). We limited these to accessible riverine wetlands adjacent to navigable rivers (to ensure federal jurisdiction) that were greater than 0.05 ha (the minimum size requirement for some assessment tools within 100 km of the Montana

**Table 1** Summary of Case Study Assessment Tools, Regulatory Need, Attributes, and Scoring

Tool	Regulatory Support	Data Attributes	Number of Functions/ Conditions	Scoring and summary scores
Montana Wetland Assessment Method ( <b>MWAM</b> ) (Berglund and McEldowney 2008)	CWA§ 404 permitting decisions	43	12	One ecological integrity summary score - the sum of all site function/value scores divided by the total maximum (11).
HGM Northern Rockies Riverine Floodplain ( <b>HGM Approach</b> ) (Hauer et al. 2002)	CWA§ 404 permitting decisions	13	8	Eight separate scores from weighted indices between 0–1 to inform wetland capacity to perform each function.
Montana Department of Environmental Quality Wetland Condition Assessment ( <b>DEQ</b> ) (MT-DEQ 2019)	CWA§ 303(d)&305(b) Assesses wetlands condition	24	5	One wetland ecological integrity summary score (0–1) from condition scores applied to a weighted index.
Montana Natural Heritage Program Montana Ecological Integrity Assessment Protocol ( <b>EIA</b> ) (MT-NHP 2018)	CWA§ 303(d)&305(b) Assesses wetlands condition	21	4	One wetland ecological integrity summary score (0–1) from condition scores applied to a weighted index.
Natural Resources Conservation Service Riparian Assessment ( <b>NRCS</b> ) (USDA 2012)	Assist landowner's conservation efforts	10	10	One ecological integrity summary score (0–1) is calculated by summing attribute scores and dividing them by the maximum score (60).
Bureau of Land Management Properly functioning conditions for lotic systems ( <b>PFC</b> ) (BLM 2015)	BLM policy H-4180-1 rangeland health	17	17	One ecological integrity summary score (0–1) is calculated by summing attribute scores and dividing them by the maximum score (19).

**Table 2** Landscape development intensity coefficients were used to calculate the LDI index (modified from Brown and Vivas 2005; Wardrop et al. 2016 to convert to a 0–1 scale). Land cover categories come from the 2016 NLCD 30 m product (MRLC 2020)

NLCD Code	NLCD Description	LDI Weights
11	Open Water	1
41	Deciduous Forest	1
42	Evergreen Forest	1
43	Mixed Forest	1
52	Shrub/Scrub	1
71	Herbaceous	1
90	Woody Wetlands	1
95	Emergent Herbaceous Wetlands	1
81	Hay/Pasture	0.63
82	Cultivated Crops	0.36
21	Developed, Open Space	0.20
22	Developed, Low Intensity	0.20
31	Barren Land	0.13
23	Developed, Medium Intensity	0
24	Developed, High Intensity	0

State University (Bozeman, MT). We assigned LDI scores for all potential wetlands buffered to 200 m (maximum landscape assessment area among the selected tools). The LDI assigns 'emergy' scores to land cover classes. We used

the 15-land cover classes in the 2016 National Land Cover Database (NLCD) 30 m product (MRLC 2020). We modified the LDI emergy scores from Brown and Vivas (2005) and converted them to a normalized 0–1 scale to facilitate comparisons with the assessment protocols, with 1 representing the lowest level of development (Table 2). To be clear, this is not how LDI is used elsewhere, but we wanted them to be comparable to the assessment tools' 0–1 score range. Because the six tools assess the condition of the interior and the buffer of the wetlands, we also calculated the emergy scores for all the land cover within the wetland and its buffer. Although the developed land classes from Table 2 will not likely occur within a wetland, it is common to have hay or pasture in Montana wetlands. These scores were then spatially averaged.

For the final site selection, the land cover within the initial buffered sites was hand-digitized using the 2018 high-resolution (0.5 m) world imagery in ArcGIS ver. 10.6.1 (ESRI 2018). Each polygon was assigned one of the NLCD 15-land cover classes, and the LDI was calculated. The final sites represent a reference domain that captures a range of wetlands found within the region based on final LDI scores and the authors' best professional judgment. Selection of reference sites using best professional judgment

by experienced practitioners is typical in assessment tool development (Brinson and Rheinhardt 1996). From these, we selected 18 sites, with LDI scores that ranged from 0.23 to 0.87, that best captured the regional scope of disturbance (Fig. 1). Lastly, we color-coded these wetlands into general categories of development based on the LDI ranges— low disturbance:  $>0.70$  (range 0.72–0.87), medium:  $>0.50$ –0.70 (range 0.53–0.68), or high disturbance:  $\leq 0.50$  (range 0.23–0.50). These course categories only illustrate each tool’s response to the gradient. The ranges were selected to establish a nearly equal number of sites in these categories.

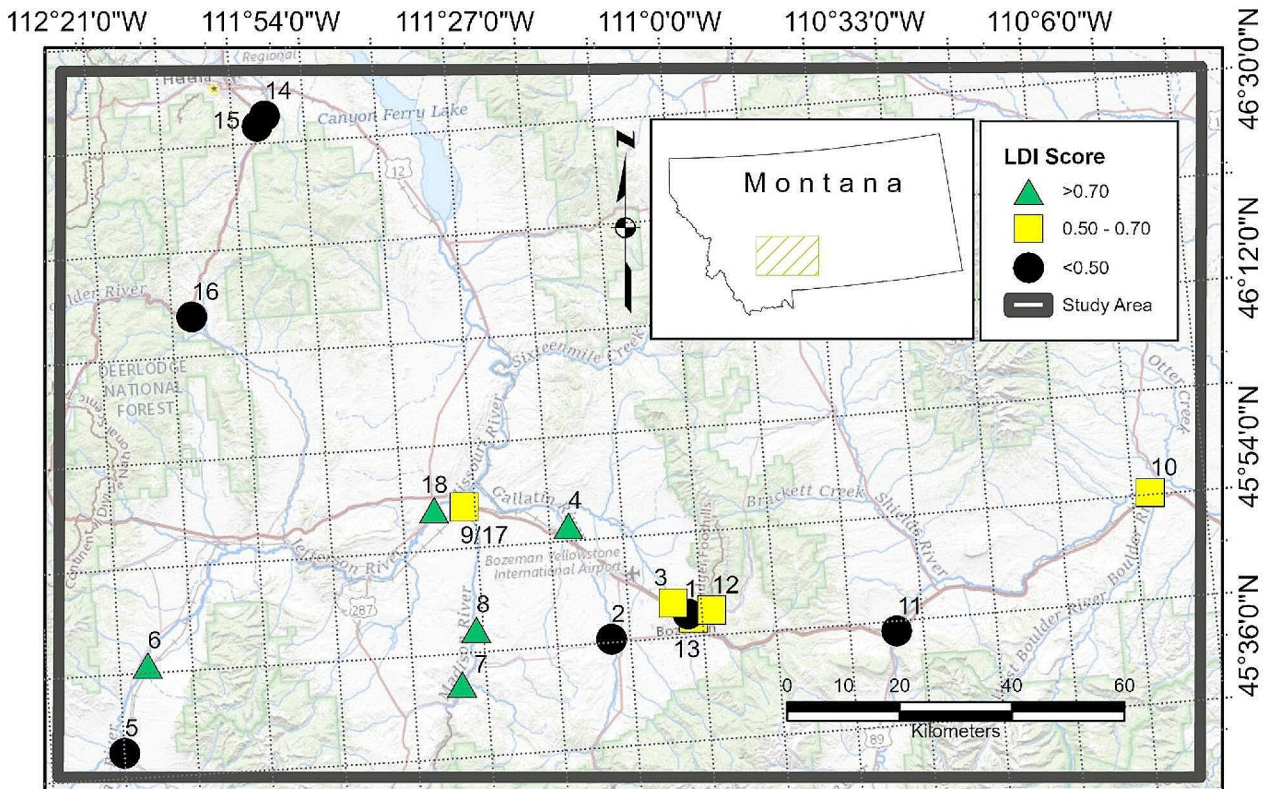
**Field Methods and Site Scoring**

The six protocols share many attributes that need quantitative measures conducted in an office (e.g., total area, relative area by cover type, the interspersion of cover) and in the field (e.g., vegetation cover and composition). We ran all six protocols simultaneously. We conducted representative 50 m line-intercept transects and 50×50 m plots within each site to account for the percent cover of herbaceous and woody species/wood debris (respectively), species composition, and native/non-native species ratios. All

protocols require the end-users to examine the entire site thoroughly for invasive species potentially missed by the established sample transects and quadrats and other qualitative observations (e.g., bank stability, browse intensity, and surface water connections). Each protocol has a different approach to the landscape assessment area (aka ‘buffer’): MWAM~ 15 m, HGM Approach– geomorphically defined, DEQ– 100 m, EIA– 200 m, NRCS– not defined, PFC– not defined. The scores for the attributes, functional capacity, and overall condition followed the procedures prescribed in each protocol.

**Analysis of Tool Performance**

This analysis, conducted in two parts, provided a logical approach to seeking those insights: Part 1 explored how well each tool captures the range of disturbance, as defined by our sample site’s LDI, through visualization, descriptive, and non-parametric statistics. Part 2 examined individual tool performance across our sample sites by assessing how each tool’s elements generalize to the concept of ecological condition as measured by their summary ecological condition scores. All data visualizations and analyses were



**Fig. 1** Overview of case study site locations in Southwestern Montana. The top right inset shows an outline of the State of Montana with a hatched box indicating the general study area. Low LDI scores in black

circles indicate more disturbed sites ( $\leq 0.50$  LDI), mid-range sites in -yellow squares ( $> 0.50$ –0.70 LDI), and least-disturbed sites in green triangles ( $> 0.70$  LDI)



conducted in the R software environment version 4.0.3 (R Core Team 2020), and the accompanying code is provided in the Supplementary Material link.

**Part One** First, to facilitate comparisons, we calculated a final condition score that ranged from 0 to 1 following each of the tool's protocols, with 1 representing the best condition. The HGM Approach does not provide a final overall score (Smith et al. 1995). However, we averaged the functional capacity index score for each guidebook's indices for comparative purposes to provide an overall score between 0 and 1. To be very clear, this is not how the HGM Approach is applied in practice. The average used here is strictly to facilitate comparison across tools. We then summarized ecological condition site scores by each tool (see Table 1) and compared them to site scores assigned by LDI through descriptive statistics. We grouped sites by LDI scores into three tiers of human impact within the wetland and buffer and assessed each tool's performance at capturing disturbance within those tiers. We conducted a Mann-Whitney U Test to compare values for each tool against the LDI scores to determine if the two groups of scores are different (with a  $H_0$  that the two populations are equal). For this test, we assume that LDI is an arbiter of disturbance, however, we are aware of the limitations of LDI. Therefore, the holistic approach served as a general guideline in assessing the ability of various tools to describe a set of sample sites, helping evaluators to decide which tool might best capture disturbance in their own applications. Lastly, we created tile graphs to visualize how each tool's metrics or indices compare to their overall score (Wei and Simko 2017; Wickham et al. 2019).

**Part Two** Because ecological condition derived from indicators is not directly measured, it is considered a latent variable. In combining the individual metric scores to create one overall site condition score, each tool assumes that all metrics contribute equally to a single underlying concept. Using an exploratory factor analysis approach, we assess the ability of each multi-metric index to truly measure ecological health.

Factor loadings were estimated with the least-squares minimum residuals (minres) algorithm using the psych package (Revelle 2019). Details on the methods and reproducible code are available in the Supplementary Material. To conduct factor analysis, tools with many functions/conditions elements relative to the sample size ( $n=18$ ) necessitated the reduction of individual function/condition measurements that informed overall site condition (MWAM and NRCS). Although the 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. For MWAM, we merged habitat metrics to form 'general habitat'; sediment and bank stability

to 'sediment source and storage'; and water storage and groundwater recharge to 'water storage exchange.' For NRCS, we merged stream incision, stream bank, and stream sediment water balance into a condition called 'geomorphic considerations.' We also merged bank vegetation, cover, noxious weeds, invasive species, and woody regeneration into a condition called 'vegetation considerations.' Additionally, factor analysis requires at least four elements to test the loading. However, the PFC model has only three indices. Therefore, we constructed an additional condition sub-index called 'carbon' based on an average metric score from flooding, riparian impairment, plant vigor, wood source, energy dissipation, and stabilizing plants.

## Utility Analysis

We surveyed wetland experts in the U.S. Intermountain West to understand which assessment tools respondents used and why. First, we asked respondents to select which of the six aquatic assessment tools they used (see Table 1). Then, we asked questions to determine respondents' level of trust in the tools' results as a means of interpretation between our determination of the tool's performance and the end-user experience. This survey was part of a larger effort, where we surveyed wetlands experts from six states ( $n=179$ ); here, we report on Montana-specific responses ( $n=56$ ). The survey asked a series of closed-ended questions, including binary response items (yes/no) and 5-point Likert scale responses (strongly disagree = 1 to strongly agree = 5).

Tool assessment questions were adapted from a survey designed to understand the efficacy of agricultural decision support tools (Ranjan et al. 2020). In December 2020, we distributed the survey through online Qualtrics survey software (Qualtrics 2020) to two groups: (1) Federal, State, Local, and Tribal wetland decision-makers (e.g., MT-DEQ, MTNHP, MT-DOT, USACE, local jurisdictions, and tribes) determined through our knowledge, publicly available data, and recommendations from other experts. We then asked those individuals to forward our request to other experts. (2) Email distribution listservs from the Montana Wetland Council, Montana Watershed Coordination Council, and the Pacific Northwest and Rocky Mountain Society of Wetland Scientists chapters. We do not know the total number of people this distribution method reached. Data were analyzed using SPSS Statistics software (IBM SPSS Statistics, Armonk, NY). The survey is covered under MSU's Institutional Review Board approval (IRB number SC100820-EX). Using survey responses, we ranked the utility of the protocols by their underlying mandates (CWA 404, 303d, BLM, or NRCS) based on how often practitioners used these tools in the last 12 months. Lastly, we provide an overall ranking of a preferred tool by their underlying mandates.

We ranked each by the highest range of assessment scores and recent use utility rank.

## Results

### Range of Conditions

#### Part One: Visualization and Summary Statistics

Our ranking of case study wetlands, based on LDI, provides a generalized glimpse into landscape drivers that affect wetland conditions (Wardrop et al. 2016). Commonly, site-specific data will reveal subtleties in a condition not captured by this approach. Table 3 displays four summary statistics (mean, minimum, maximum, and range) for each assessment tool's overall site scores, starting with the LDI as a reference. The HGM Approach shows the highest range of scores (0.72) and the lowest minimum score (0.19). The DEQ and EIA have the most limited ranges (0.28 and 0.29, respectively) and the highest minimum site scores (0.63 and 0.61, respectively). The PFC and NRCS have the highest mean site scores, both at 0.86, giving some sites a perfect score of one, indicating some locations are in an ideal ecological condition. The Mann-Whitney U Test provides strong evidence that the distribution of the EIA, NRCS, and BLM scores differs significantly from the LDI scores. On the contrary, the higher p-values of MWAM, HGM Approach, and DEQ suggest a strong relationship with the distribution of disturbance measured by LDI. More detailed data on the distribution of each tool is available in the supplemental material.

Fig. 2 represents how each of the six assessment protocol summary scores distributes across the disturbance gradient across the sample. Scores are grouped based on the LDI scale: greater than 0.70 are considered 'less disturbed'

**Table 3** Summary statistics across sites by assessment tool. Average overall score, minimum overall site score, maximum overall site score, and observed range of overall site scores for the sample

Assessment Tool	Mean Score <sup>1</sup>	Min Score	Max Score	Score Range	Mann-Whitney U Test <sup>2</sup> P-value
LDI	0.58	0.23	0.87	0.64	-
MWAM	0.56	0.35	0.75	0.40	0.76
HGM Approach	0.61	0.19	0.91	0.72	0.60
DEQ	0.77	0.63	0.91	0.28	0.20
EIA	0.78	0.61	0.90	0.29	<0.01
NRCS	0.86	0.43	1.00	0.57	<0.01
PFC	0.86	0.44	1.00	0.56	<0.01

<sup>1</sup>LDI is a development score, and HGM is a functional capacity score rather than a condition score.

<sup>2</sup>Test compares tool scores to LDI scores.

(green triangles), scores between 0.50 and 0.70 are 'mid-range' sites (yellow squares), and those scoring less than 0.50 are 'more disturbed' (black circles) sites (see Fig. 1). Average site scores assigned by each tool, indicating overall ecological condition, are displayed in Fig. 2. Tool names appear on the y-axis, and the average site score on the x-axis.

Tile plots are a helpful way of visualizing high-dimensional data. Fig. 3 displays tile plots for each tool to visually compare scoring patterns within each index over the range of sites to see overall site score patterns. Multi-metric indices and the overall score by site appear on the y-axis. The score is on the x-axis, arranged in descending order of each tool's overall site score from left to right. Coloring indicates scores, 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' (called 'average' in Fig. 3).

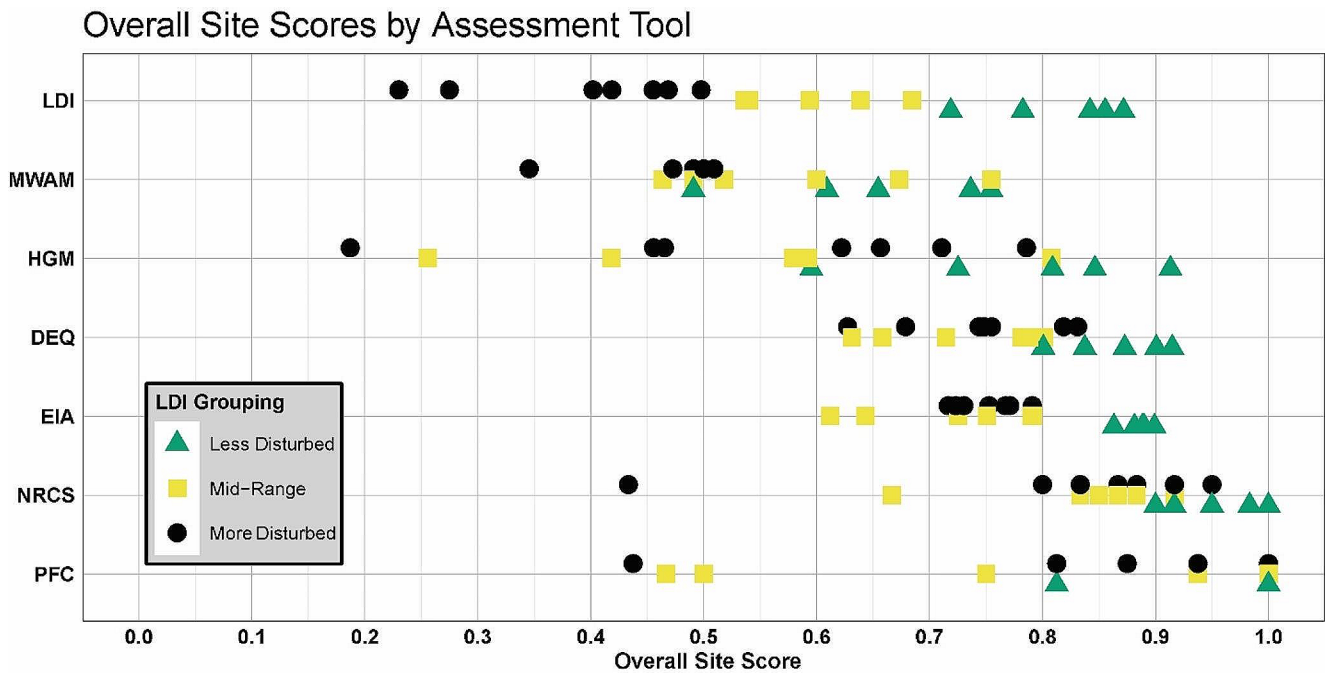
#### Part Two: Factor Analysis

Factor loadings measure the degree of generalizability and reflect the quantitative relationships between each function/condition measurement (assessment index) and a factor (overall site condition as measured by the tool's summary score). These tools combine multiple indices (or metrics) into a single factor, 'summary ecological condition score.' The farther the factor loading is from zero, the more one can generalize that the assessment index relates to the overall site condition (Gorsuch 1983). Results of factor analysis, displayed in Fig. 4, show the numerical estimates of factor loadings for each tool. Negative factor loadings are in shades of red, and positive are in shades of blue, with colors closer to white indicating a weaker loading. Weak or negative factor loadings suggest that some individual condition indices/metrics do not contribute to the concept of wetland ecological condition in the way that the tool assumes. See supplemental material for details.

### Tool Utility

Our social science survey indicates which tools respondents use and provides information about perceptions of tool utility. Here, we report on responses from Montana, who answered: "yes" to the question, "Do you, or have you ever, used an aquatic ecosystems assessment tool?" ( $n=56$ ). Note that not all respondents answered all questions; thus, response numbers may not equal the total number who responded to the assessment tool's filter questions.

We asked respondents to indicate which aquatic assessment tools they have used in Montana and found that the top three were PFC ( $n=23$ ), DEQ ( $n=21$ ), and MWAM



**Fig. 2** Summary site scores for the entire sample by each assessment tool, including LDI criteria. Tool names appear on the y-axis, and site condition scores on the x-axis. Individual points indicate each site's ecological condition summary score (see Table 1). Data points are ver-

tically jittered 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

( $n=20$ ). We then asked respondents to select which tool they used the most in the previous 12 months and found that DEQ ( $n=11$ ) and MWAM ( $n=11$ ) were the top two tools, closely followed by PFC ( $n=7$ ) and EIA ( $n=6$ ). Few respondents selected NRCS ( $n=3$ ) or HGM ( $n=1$ ) (Table 4). Lastly, we ranked utility by the extent of its most recent use.

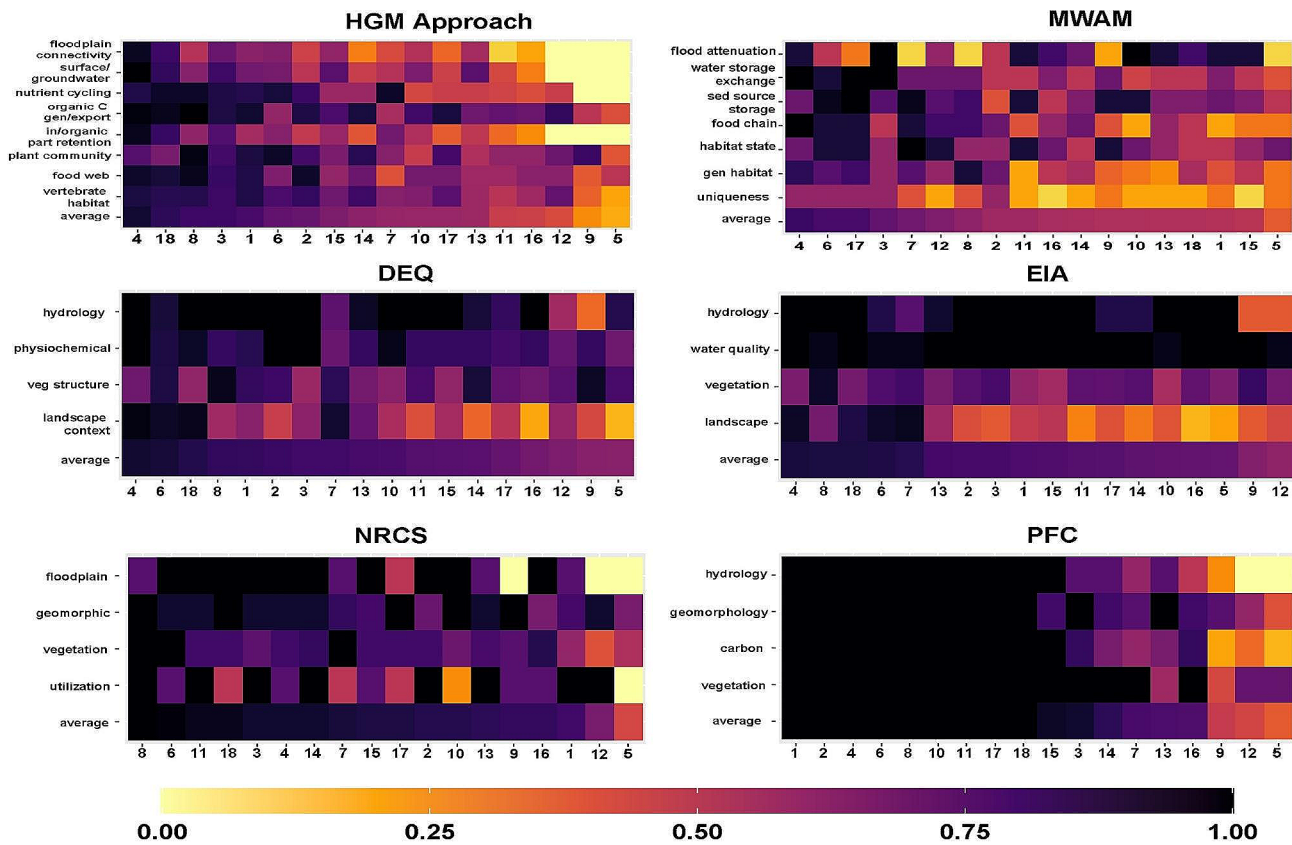
We then asked respondents to rate seven statements about their trust of their most used tool by agreement or disagreement with the statement (strongly disagree = 1 to strongly agree = 5) (Table 5). Each statement shows various response numbers depending on how respondents chose to answer. We report respondents' perceptions of the tools to inform current attitudes toward each tool and aid in future tool development.

Eleven people reported using the MWAM tool the most. Of these, users felt slightly under the "agree" for the statement "*I trust the data produced by the tool,*" indicating general trust in the data produced by MWAM (mean = 3.90;  $n=10$ ). Users agreed that MWAM has replicable results (mean = 4.00;  $n=10$ ). Respondents were close to neutral in their agreement that MWAM is sensitive enough to detect impacts from various stressors (mean = 3.20;  $n=10$ ). There was also less agreement that the MWAM tool adequately captures the importance of various functions at assessment locations (mean = 3.60;  $n=10$ ) and adequately distinguishes across a range of conditions (mean = 3.50;  $n=10$ ).

Only one individual used the HGM Approach, and they agreed (mean = 4.00) that it captures the importance of various functions at assessment locations and strongly agreed (mean = 5.00) that the tool could distinguish across a range.

Eleven people reported using the DEQ, and respondents agreed that the DEQ tool is trustworthy (mean = 4.25;  $n=8$ ), produces replicable results (mean = 4.00;  $n=7$ ), and is sensitive enough to detect impacts from various stressors, and captures the importance of various functions (mean = 4.00;  $n=7$ ). Respondents had less agreement (although this mean was very close to "agree") that the DEQ tool adequately distinguishes across a range of conditions (mean = 3.86;  $n=7$ ). Six people reported using the EIA, and respondents agreed that the EIA tools are trustworthy (mean = 4.00;  $n=2$ ), produce replicable results (mean = 4.15;  $n=2$ ), are sensitive enough to detect impacts from various stressors, and capture the importance of various functions (mean = 4.00;  $n=2$ ). Respondents also agreed that the EIA tool adequately distinguishes across a range of conditions (mean = 4.00;  $n=2$ ).

Seven people reported that they used the PFC tool the most. The PFC tool had low ratings for trustworthiness of the output (mean = 3.60;  $n=5$ ), data produced (mean = 3.80;  $n=5$ ), and replicability of the results (mean = 2.80;  $n=5$ ). The respondents also rated the PFC tool low for its ability to detect impacts from various stressors (mean = 2.80;  $n=5$ ) and to adequately distinguish across a range of conditions (mean = 3.20;  $n=5$ ). Respondents had less agreement



**Fig. 3** Tile plot of index scores and overall tool score by site for each tool. Indices and overall average scores are listed on the y-axis, with sites on the x-axis. From the left to the right-hand side of the x-axis,

sites are arranged in descending order of the overall tool score. Coloring indicates score, with lower scores in lighter colors and higher scores in darker colors

that the PFC tool could capture the importance of various functions (mean=3.40;  $n=5$ ). Three people reported that they used the NRCS tool, and overall, NRCS rated higher than PFC tools from the limited respondents for the trustworthiness of the output (mean=4.20;  $n=1$ ), data produced (mean=4.20;  $n=2$ ), and replicability of the results (mean=4.00;  $n=2$ ), ability to detect impacts from various stressors (mean=4.00;  $n=2$ ) and to adequately distinguish across a range of conditions (mean=4.00;  $n=2$ ). Respondents had less agreement that the tools could capture the importance of various functions (mean=3.50;  $n=2$ ).

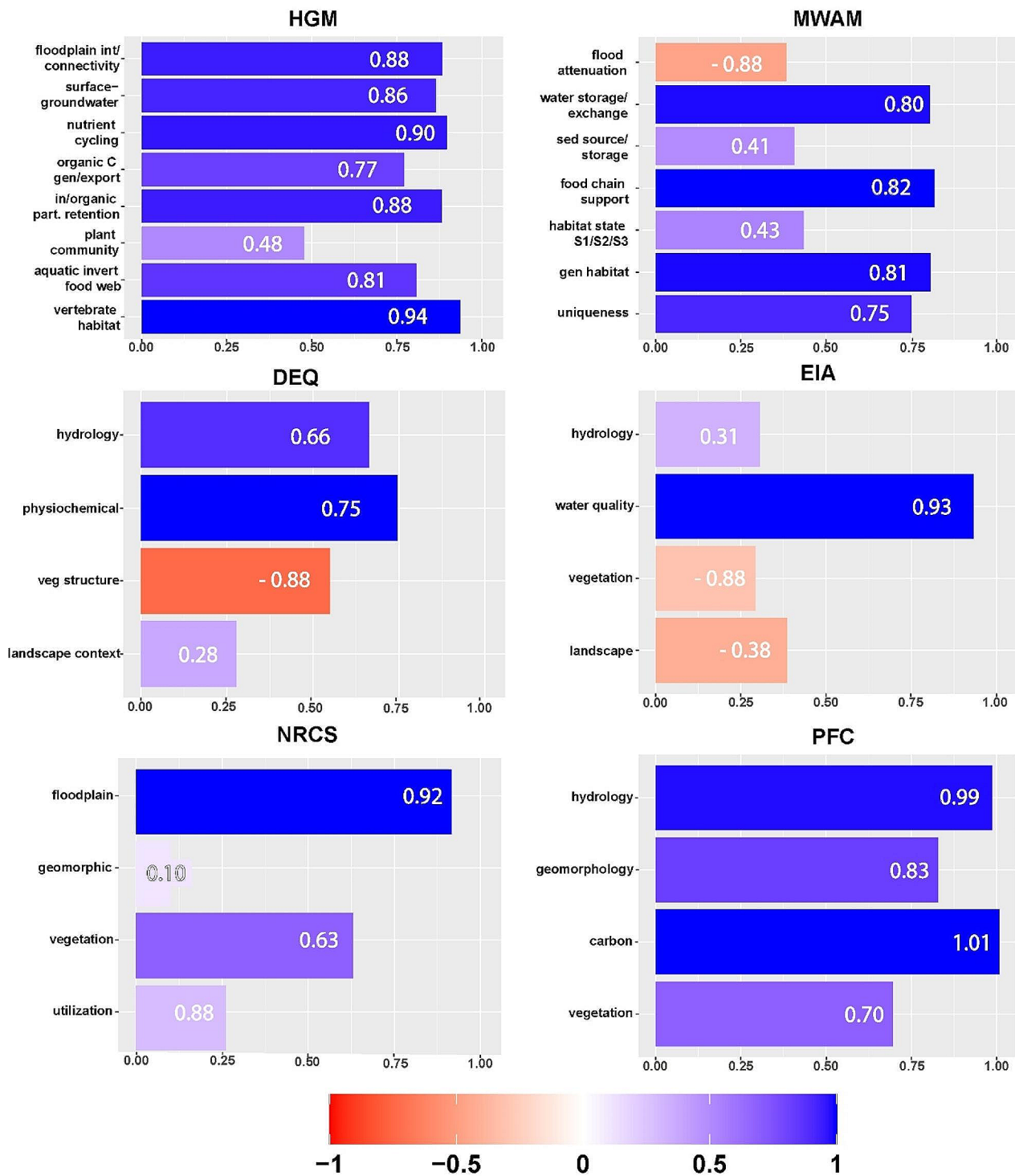
### Discussion

One would expect a general agreement across wetland assessment tools based on the concept that increased human alterations are a primary stressor that diverges wetland ecosystems from an ideal state (Rapport et al. 1985; Fennessy et al. 2007). Here, we use LDI as a surrogate for multiple human disturbance drivers. Yet, we recognize that this coarse assessment is not a comprehensive measure of more subtle disturbance elements found within a wetland.

However, LDI provides a general distribution of wetlands across a disturbance gradient. If it was the intent of each assessment approach to document departure from an ideal, least disturbed state (viz. Stein et al. 2009), then several of the tools tested are not operating as the developers hoped. We found that many metrics that make up each tool do not capture the same distribution of disturbance as other metrics within the same tool and are not well integrated into an overall condition score. Once again, this may result in the tool operating in a manner not intended by the developer. If these tools are not acting as intended, this may have ramifications on determining the extent of unavoidable impacts and resulting compensatory mitigation. They may not accurately determine the condition of the aquatic resources being reported to Congress. Or, if the assessment tool helps make required management decisions, such as closing grazing allotments, the resource may be far more damaged than that decision threshold.

We examined each tool’s overall performance relative to the LDI scores. Our final sites do not represent pristine areas devoid of disturbance nor completely impacted sites as one might find in a highly urbanized system. Our final site selection intended to reflect a disturbance gradient typical for





**Fig. 4** Factor analysis results for each tool are displayed in a bar chart, along with numerical estimates of factor loadings. The absolute value of the estimated factor loadings is plotted for easier comparison

between index loadings for each tool. Negative factor loadings are in shades of red, and positive factor loadings are in shades of blue, with colors closer to white indicating a weaker loading

Montana’s riparian wetlands as defined by the LDI, recognizing that our selected sites are not the final arbiter of that disturbance gradient. We assume four things are true about each assessment tool and our selected sites. First, our sites

represent a distribution of regional wetlands ranging from ecologically intact to disturbed. Second, ecologically intact wetlands have higher ecological conditions than disturbed wetlands. Third, these assessment tools provide some

**Table 4** Respondent's response to the question, "You specified that you primarily work in Montana. Please indicate whether or not you use the following aquatic system assessment tools in that state". Followed by the response to, "You specified that you use the aquatic system assessment tools listed below. Please select the tool you used the most in the last 12 months." ( $n=38$ )

Tool	Yes (use the tool in MT)	No (do not use the tool in MT)	The tool used most in the last 12 months	Rank by Recent Use
MWAM ( $n=37$ )	20	17	11	1
HGM Approach ( $n=37$ )	12	25	1	5
DEQ ( $n=40$ )	21	19	11	1
EIA ( $n=39$ )	16	23	6	3
NRCS ( $n=39$ )	17	22	3	2
PFC ( $n=33$ )	23	10	7	4

reliable measurement of that ecological condition. Fourth, we recognize that each tool measures ecological conditions differently. Yet, we assume that the tool would also distribute the sites across a gradient, although perhaps in a different order than those selected by our iterative approach and represented by the wetland's LDI score. This assumption is supported by Mack (2006). The rapid assessment tools tested here incorporate site-specific details on ecological processes or conditions that go beyond the land use gradient of the LDI. As a result, it is not surprising that the six tools do not align directly with the LDI scores. It should be noted that one element we did not address is user consistency and repeatability. When measurements of assessment variables are not repeatable, the conditions scores are detrimentally affected, especially functions modeled by only a few variables. See Whigham et al. (1999) and Herlihy et al. (2009) for more details on this crucial aspect.

For this study, we also incorporated social science data to understand the users' use of and trust in each tool. These data were combined with the LDI results related to the overall performance of each tool. We recognize there are limitations to interpretations of the social science data, particularly concerning low respondent numbers for the HGM and NRCS tools. However, the limited number of responses

tell their own story: some tools are not used as often as others, but they show promise in their utility. This could warrant more exploration to understand why the tools are not used or how the utility in those tools could be transferred to more used tools.

### Support for CWAS 404 (MWAM and HGM Approach)

The U.S. 2008 Compensatory Mitigation Rule (CMR) puts the burden on the U.S. Army Corps of Engineers to determine if compensatory mitigation is sufficient to replace lost aquatic resource functions (33 CFR at § 332.3(f)(1); G.P.O. 2008). It follows, therefore, that a tool that best reflects the wetland's underlying condition would best assist in proper compensatory mitigation. The field study results found that HGM Approach's summary scores had the widest distribution with a range of 0.72 (0.19–0.91) and were very similar to the LDI distribution (Mann-Whitney  $p$ -value of 0.60), and MWAM was more confined to the middle of the LDI distribution with a range of 0.40 (0.35–0.75). Yet, it was the most similar to the LDI distribution (Mann-Whitney  $p$ -value of 0.76). However, many MWAM scores are concentrated toward the middle of the distribution (See Fig. 2). This would have mitigation implications as the tool may undercompensate for impacts to higher-quality sites and overcompensate for lower-quality locations. This may be what the users recognize as well. It is the preferred tool with Montana's highest utility wetland permitting process (see Table 4). Yet respondents fell between 'neutral' and 'agree' when asked if MWAM adequately distinguishes across a range of conditions (mean=3.50;  $n=10$ ). MWAM is designed to apply across the broader physiographic regions and multiple wetland classes in Montana, relying on professional experience and literature rather than reference data to establish sub-index scores (Berglund and McEldowney 2008). This could also account for individual MWAM metrics or indices providing information that contradicts the averaged condition score indicated in our factor analysis.

The MWAM tile graphs and factor analysis (See Figs. 3 and 4) results suggest that many but not all individual

**Table 5** Respondents' level of disagreement or agreement with the following statements regarding the aquatic assessment tool they used most often for aquatic system assessment in the past 12 months (strongly disagree = 1 to strongly agree = 5) followed by overall utility rank

Statement	Mean and number (in parentheses) of responses by the tool					
	MWAM	HGM	DEQ	EIA	NRCS	PFC
The stakeholders I work with trust the outputs of the tool	4.33 (9)	4.00 (1)	4.29 (7)	4.00 (1)	4.00 (1)	3.60 (5)
I trust the data produced by the tool	3.90 (10)	5.00 (1)	4.25 (8)	4.00 (2)	4.00 (2)	3.80 (5)
Produces replicable results	4.00 (10)	5.00 (1)	4.00 (7)	4.50 (2)	4.00 (2)	2.80 (5)
Sensitive enough to detect impacts from various stressors	3.20 (10)	4.00 (1)	4.00 (7)	4.00 (2)	4.00 (2)	2.80 (5)
Adequately captures the importance of various functions at assessment locations	3.60 (10)	4.00 (1)	4.00 (7)	4.00 (2)	3.50 (2)	3.40 (5)
Adequately distinguishes across a range of conditions	3.50 (10)	5.00 (1)	3.86 (7)	4.00 (2)	4.00 (2)	3.20 (5)
<b>Average mean of all statements</b>	<b>3.88</b>	<b>4.57</b>	<b>4.08</b>	<b>4.00</b>	<b>3.93</b>	<b>3.34</b>

function/condition measures generalize well to the single latent variable of overall wetland condition. For example, in Fig. 3, we can see that many sites that scored high for flood attenuation scored, yet very low for Uniqueness and vice-versa. Flood attenuation is a measurement designed to combine roughness and opportunity to receive floodwaters, and many wetlands have similar scores despite, or contradictory, to other indicators of disturbance. Because this measurement has equal weight with all other assessment indices when determining the overall MWAM condition score, it may raise the scores of impacted wetlands and lower healthy wetlands scores. Uniqueness is designed to capture rare wetland types (i.e., fens or bogs) or uncommon wetlands that are also structurally diverse. Because riparian wetlands are not rare and relatively abundant in the study area, obtaining a high score for this metric is difficult. Therefore, many wetlands have similar scores despite, or contradictory, to other indicators of disturbance (see Fig. 2). In Fig. 4, the factor analysis indicates that metrics like Uniqueness strongly influence the overall score of the wetland condition (the latent variable), as all scored relatively the same across the gradient. However, the distribution of Flood Attenuation scores does not match the range of overall condition from the average of all metrics and is not well integrated into that latent variable. Similar arguments can be made for sediment sources and state-listed species. Ultimately, this tool may not operate as assumed as it is integrated into an overall condition score for mitigation requirements.

The HGM Approach is designed to assess riverine wetlands in the Northern Rockies and is the only tool among the six tested that uses reference conditions to scale their metrics. This could account for the tool capturing the widest score range across the selected disturbance gradient. Although the model is more sensitive to disturbance gradients than the others in the study, it is not the user preferred tool for the wetland permitting process. However, the Tile Graph and Factor Analysis figures show that most indices align with the overall score, with a minor exception of Plant Community. This index measures vegetation composition and could score higher despite lower floodplain connectivity or inorganic particle retention scores. As stated earlier, the HGM Approach was designed so that each function is treated separately and has a history of resisting averaging the scores into a single variable. This is the problem described in the MWAM above, where averaging the indices to a single score reduces the sensitivity of individual models within the tool. However, single scores are often preferred in debit/credit determination to account for mitigation success and permit efficiency (Lave and Doyle 2020).

### Support for CWA§ 303(d) & 305(b) and State Monitoring (DEQ and EIA)

State wetland monitoring and assessment programs help communities develop and implement watershed plans to meet water quality standards and protect aquatic resources. These programs allow states to establish baseline conditions, detect change, and characterize trends in the conditions of aquatic resources (USEPA 2015). In Montana, both the Department of Environmental Quality (MT-DEQ) and Natural Heritage Program (MT-NHP) assessed wetland and riparian sites to establish a statewide reference network to help inform wetland and riparian resource management, planning decisions and restoration efforts (MT-DEQ 2013). Once again, a tool that best reflects the wetland's underlying condition would assist in this effort. MT-NHP initially developed a multi-level assessment approach that includes mapping, a rapid assessment tool (the EIA), and highly detailed data gathering (MT-NHP 2018). DEQ modified these approaches for their three-tiered approach (MT-DEQ 2019). Both rapid tools are not based on these reference data but on professional experience and literature to establish sub-index scores.

We found that both tools' overall condition scores compress the sites toward a higher condition than was reflected in our LDI or the HGM/MWAM approaches. DEQ has a range of 0.28 (from 0.63 to 0.91) and is similar to the LDI distribution (Mann-Whitney p-value of 0.20). EIA has a range of 0.29 (from 0.61 to 0.90), but with strong evidence that the distribution of scores is different than the LDI distribution (Mann-Whitney p-value of <0.01). Neither tool can discriminate the most impacted sites from the others in the gradient within our sample sites (See Fig. 2). Yet, the users of these tools generally agreed that the tools could distinguish across a range of conditions and detect a variety of stressors, although the sample size of the users is small (DEQ  $n=7$  and EIA  $n=2$ ).

Despite their common foundations, these tools behave differently. For instance, the EIA summary condition score has a strong estimated loading only for the water quality index (0.93). The Tile Graph shows that this metric scored very high for each site, resulting in a higher overall score across sites. This tool bases water quality assessment on visual evidence of water clarity, oil sheen, and abundance of eutrophic species. Although these are excellent indicators of poor water quality, they are rare except for highly disturbed wetlands. The DEQ and EIA, vegetation index scores, rely on a combination of native plant cover, aggressive graminoids, noxious weeds cover, herbaceous litter, regeneration affected by litter, wood regeneration, and herbivory. Many of these harmful elements, such as high ungulate populations, aggressive grasses, and noxious weeds, are omnipresent in

riverine wetlands in Montana and are an indication of watershed scale disturbance and not as reflective of a local disturbance gradient as some of the physicochemical attributes. DEQ has a wider scoring range than many of the vegetation attributes than the EIA tool and, therefore, more sensitivity to a score range for the index. The Tile Graph shows that many sites had higher vegetation structure scores than the overall DEQ tool range, leading to a negative loading (-0.55) on the condition score. In both tools, weak or negative factor loadings suggest that some of the individual condition metrics/sub-indices do not contribute to the concept of wetland ecological condition in the way the tool assumes.

Both tools do a poor job of distinguishing sites across the disturbance gradient. However, the DEQ tool had the highest utility of the two tools, given its higher use (see Table 4). Therefore, while aware of the tool's shortcomings, we recommend DEQ as the preferred tool.

### **Assist Aquatic Landowners and BLM Policy H-4180-1 Rangeland Health Standards (NRCS and PFC)**

Under directives in the USDA National Planning Procedures (USDA 2014), the NRCS uses its riparian assessment tool to prioritize and direct resources to prevent further degradation and achieve the greatest return for the investment. Under Title 43, Sect. 4180 of the Code of Federal Regulations, the BLM must provide measures and guidelines to improve the health of public rangelands (BLM 2001). Their lotic PFC models assist in riparian aquatic systems in these rangelands (BLM 2015). The NRCS tool was influenced by elements of the BLM lotic PFC model to help measure riparian systems' stability and sustainability (USDA 2012). Both rapid tools are not based on these reference data but rather on professional experience and literature to establish metric scores. Both tools are designed for monitoring and assessment to help with immediate management decisions of lotic systems and their adjacent riparian areas. All our assessment wetlands have lotic shorelines and riparian areas, so these tools were appropriate. Understandably, for a landowner interested in seeing if their land practices (NRCS tool) or if grazing on BLM lands (PFC) are doing severe or immediate harm to riparian wetlands, these coarse assessment tools may be appropriate. Yet, these tools may not capture long-term trends on the ecological processes within these systems as the DEQ/EIA, nor assist with the permitting process as the HGM/MWAM approaches.

The NRCS tool has one of the 18 sites score a perfect 1.00 and seven more above 0.90 with an overall range of 0.57 (from 0.43 to 1.00). The PFC has nine sites that score an ideal 1.00 with a range of 0.56 (from 0.44 to 1.00). Both have the highest mean score across all tools (0.86 for each), and there is strong evidence that the distribution of scores

for both tools is different than the LDI distribution (Mann-Whitney  $p$ -value of  $<0.01$ ).

The metric scores within these tools are strongly influenced by energy distribution from floods, healthy plant communities with stable binding root systems to prevent erosion, and stable geomorphic considerations. The PFC has binary scores with 'one' indicating that the metric is present and broadly provides these criteria. To achieve a score of 'zero,' there needs to be an overt indication that the metric is not meeting the criteria. PFC separated some disturbed sites consistently across the metrics, as seen in the Tile Graphs, and the Factor Analysis indicates that the tool is acting as intended. However, most sites scored a 'one' for many metrics, and the tool lacked the sensitivity to distinguish between the established disturbance gradients.

The NRCS tool provides a range of scores for their metrics, yet these metrics do not adequately capture the gradient of disturbance. For instance, in Montana's gravel and cobble-dominated floodplains, where many riverine wetlands are located, there are also active shifting mosaics floodplain habitats driven by fluvial processes that cause eroding banks (Stanford et al. 2005; Kleindl et al. 2015). These are found in what is considered healthy ecosystems (Hauer et al. 2002). Yet these would score low, while modified banks that prevent erosion would score high in this protocol. The Tile Graphs and the Factor Analysis show that this metric did not generalize well to ecological conditions. In a second example, the NRCS tool, like DEQ and EIA, has metrics that account for ungulate utilization through indications of browse. Ungulates are widespread in Montana and are not well related to general disturbance gradients as they occur in the best and worst sites. The Tile Graphs and the Factor Analysis show that this metric did not generalize well to ecological conditions. The mixed results of the NRCS tool indicate that not all indices generalize well to ecological conditions and that the model may not function as assumed for this sample.

The NRCS tool users agreed that they could detect impacts from stressors and that the tool was repeatable, yet the PFC users responded less than neutral to these statements (see Table 5). Yet, both tools do a poor job of distinguishing sites across the disturbance gradient. However, each tool provides a different objectives; therefore, while being aware of each tool's shortcomings, NRCS and PFC are equally preferred to meet the end-user's needs. Yet, if the tool is not operating as intended, these immediate management decisions may not be based on the best available information.



## Conclusions

Ecological assessment has a long history of providing decision-makers with straightforward analysis to facilitate translating ecological knowledge to meet the management mandates of state and federal aquatic monitoring, assessment, and permitting programs (Barbour et al. 1999). This facilitation, not a search for statistical relationships or significance, drives the design and analytics of many assessment tools (Karr and Chu 1998). However, this straightforward facilitation also makes it troublesome to compare tools empirically. We acknowledge that meeting the needs of these agencies is paramount in the tool's development and use. Yet, many of these assessment tools may not operate as the developers hoped. This exercise provides a programmatic approach to compare these tools and look for weaknesses within individual tools. Within an agency, comparing different assessment tools may not be helpful. Still, a close look at the internal workings of the tool through factor analysis and tile graphs may improve the ability of the tool to meet the organization's needs. Those improvements may benefit an end user's management goals and are worth examination.

Overall, we recommend the following tools for use in Montana: to meet the needs of CWA§ 404, we recommend MWAM and the HGM Approach. While MWAM has the highest utility as reported by the survey respondents, it appears that the HGM Approach performed equally (and likely better) in the comparative analysis by displaying the highest degree of dexterity and may be a better tool to meet the goals of no-net-loss in compensatory mitigation. For CWA§ 303(d) & 305(b), we recommend DEQ, and for Aquatic landowners and BLM for policy H-4180-1 range-land health standards: NRCS and PFC, respectively. This is an observational study; therefore, no causal inference can be extended to a larger area beyond the sample sites. Please see the supplemental material for further discussion on the limitations of this study.

**Acknowledgements** We would like to thank Brianna Whitehead for her field contributions, Dr. Ric Hauer for his courtesy review, and the journals reviewer for their excellent comments. We also used ChatGPT to generate a 'snappy' title from the text in the abstract.

**Author Contributions** WJK conceived, directed, and coordinated the theoretical foundations of the work with substantial feedback from all authors; WJK led the writing of the paper with assistance from all co-authors (SPC, MCR, and RU). All authors (WJK, SPC, MCR, and RU) contributed to the study's conception and design. WJK, SPC, and RU performed material preparation, data collection, and analysis. WJK wrote the first draft of the manuscript, and all authors (SPC, MCR, and RU) commented and contributed to subsequent versions. All authors (WJK, SPC, MCR, and RU) read and approved the final manuscript.

**Funding** We acknowledge funding support from the National Science Foundation (NSF) Emerging Frontiers Macrosystems Biology program grants 1702029 and the U.S. Geological Survey Water Center

Support (USGS 104b) to Montana State University. This research was also partially supported by the Initiative for Regulation and Applied Economic Analysis and the Montana Water Center at Montana State University. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

**Data availability** The datasets generated during and analyzed during the current study are available from the corresponding author on reasonable request and in the supplemental material in the following link: <https://scholarworks.montana.edu/xmlui/handle/1/17770>.

## Declarations

**Competing Interests** Mark Rains is currently serving as the Chief Science Officer for the State of Florida. The views expressed in this article are his own and do not necessarily reflect the views of the State of Florida. Rachel Ulrich is an employee of the Los Alamos National Laboratory (LANL), and this article was reviewed and approved by LANL (LA-UR-23-24599). The authors declare no conflict of interest. The founding sponsors had no role in the study's design, in the collection, analysis, or interpretation of data, in the writing of the manuscript, or in the decision to publish the results.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## Literature Cited

- Barbour MT, Gerritsen J, Snyder BD, Stribling JB (1999) Rapid Bio-assessment Protocols for Use in Streams and Wadeable Rivers: Periphyton, Benthic Macroinvertebrates and Fish. Second Edition. U.S. Environmental Protection Agency; Office of Water, Washington, D.C
- Berglund J, McEldowney R (2008) MDT Montana wetland assessment method. 42
- Bezombes L, Gaucherand S, Kerbirou C, et al. (2017) Ecological equivalence Assessment methods: what Trade-Offs between Operationality. *Sci Basis Comprehensiveness? Environ Manage* 60:216–230. <https://doi.org/10.1007/s00267-017-0877-5>
- Bick H (1963) A review of central European methods for the biological estimation of water pollution levels. *Bull World Health Organ* 29:401
- Birk S, Bonne W, Borja A, et al. (2012) Three hundred ways to assess Europe's surface waters: an almost complete overview of biological methods to implement the Water Framework Directive. *Ecol Ind* 18:31–41. <https://doi.org/10.1016/j.ecolind.2011.10.009>
- BLM (2015) Riparian area management: proper functioning condition assessment for lotic areas. Bureau of Land Management, National Operations Center, Denver, Colorado

- Brinson MM, Rheinhardt R (1996) The role of reference wetlands in functional assessment and mitigation. *Ecol Appl [ECOL APPL]* 6:69–76
- Brown MT, Vivas MB (2005) Landscape development intensity index. *Environ Monit Assess* 101:289–309
- Cairns J Jr, Pratt JR (1993) A history of biological monitoring using benthic macroinvertebrates. *Freshw Biomonitoring Benthic Macroinvertebrates* 10:27
- Fennessy MS, Jacobs AD, Kentula ME (2004) Review of rapid methods for assessing wetland condition. U.S. Environmental Protection Agency, Washington, D.C.
- Fennessy MS, Jacobs AD, Kentula ME (2007) An evaluation of rapid methods for assessing the ecological condition of wetlands. *Wetlands* 27:543–560
- Gaucherand S, Schwoertzig E, Clement J-C, et al. (2015) The Cultural dimensions of Freshwater Wetland assessments: lessons learned from the application of US Rapid Assessment Methods in France. *Environ Manage* 56:245–259. <https://doi.org/10.1007/s00267-015-0487-z>
- Goodrich C, Huggins DG, Everhart RC, Smith EF (2005) Summary of state and national biological and physical habitat assessment methods with a focus on US EPA region 7 states. 87
- Gorsuch RL (1983) Factor analysis, 2nd edn. L. Erlbaum Associates, Hillsdale, N.J
- Hauer FR, Cook BJ, Gilbert MC, et al. (2002) A regional guidebook for applying the hydrogeomorphic approach to assessing wetland functions of riverine floodplains in the northern Rocky Mountains. U.S. Army Engineer Research and Development Center, Vicksburg, MS
- Herlihy AT, Sifneos J, Bason C, et al. (2009) An Approach for evaluating the repeatability of Rapid Wetland Assessment methods: the effects of Training and Experience. *Environ Manage* 44:369–377. <https://doi.org/10.1007/s00267-009-9316-6>
- Hynes HBN (1974) The biology of polluted waters. University of Toronto, Toronto, Ont
- Karr JR, Chu EW (1998) Restoring life in running waters: better biological monitoring. Island, Washington, D.C.
- Kleindl WJ, Rains MC, Hauer FR (2010) HGM is a rapid assessment. Clearing the confusion
- Kleindl WJ, Rains MC, Marshall LA, Hauer FR (2015) Fire and flood expand the floodplain shifting habitat mosaic concept. *Freshw Sci* 34:1366–1382. <https://doi.org/10.1086/684016>
- Kolkwitz R, Marsson M (1908) Ökologie Der Pflanzlichen Saprobien. *Ber Dtsch Bot Ges* 26:505–519
- Kusler J (2006) Recommendations for reconciling wetland assessment techniques. 130
- Lave R, Doyle M (2020) Streams of revenue: the restoration economy and the ecosystems it creates. The MIT, Cambridge, Massachusetts
- Mack JJ (2006) Landscape as a predictor of wetland condition: an evaluation of the Landscape Development Index (LDI) with a large reference wetland dataset from Ohio. *Environ Monit Assess* 120:221–241
- MRLC (Multi-Resolution Land Characteristics Consortium) (2020) National Land Cover Database. In: National Land Cover Database (NLCD). <http://www.mrlc.gov/index.php>. Accessed 1 Apr 2020
- MT-DEQ (2019) Montana DEQ Wetland Program: Wetland Assessment Protocol - Field Manual (Modified from the MTNHP EIA Protocol 2015). 92
- MT-NHP (Montana Natural Heritage Program) (2018) Montana Ecological Integrity Assessment Field Manual. Montana Natural Heritage Program, Helena, MT
- Poikane S, Salas Herrero F, Kelly MG, et al. (2020) European aquatic ecological assessment methods: a critical review of their sensitivity to key pressures. *Sci Total Environ* 740:140075. <https://doi.org/10.1016/j.scitotenv.2020.140075>
- Qualtrics (2020) Qualtrics Survey Software
- R Core Team (2020) R: a Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria
- Ranjan P, Duriancik LF, Moriasi DN, et al. (2020) Understanding the use of decision support tools by conservation professionals and their education and training needs: an application of the reasoned Action Approach. *J Soil Water Conserv* 75:387–399. <https://doi.org/10.2489/jswc.75.3.387>
- Rapport DJ, Regier HA, Hutchinson TC (1985) Ecosystem Behavior under stress. *Am Nat* 125:617–640. <https://doi.org/10.1086/284368>
- Revelle WR (2019) psych: Procedures for Personality and Psychological Research
- SCOTUS (Supreme Court of the U.S.) (1892) Illinois Central R. Co. v. Illinois
- SCOTUS (Supreme Court of the U.S.) (1950) United States v. Gerlach Live Stock Co
- Smith RD, Ammann A, Bartoldus C, Brinson MM (1995) Approach for assessing Wetland functions using Hydrogeomorphic classification, reference wetlands, and functional indices. U.S. Army Engineer Waterways Experiment Station, Vicksburg, MS
- Stanford JA, Lorang MS, Hauer FR (2005) The shifting habitat mosaic of river ecosystems. *Verhandlungen Des Internationalen Verein Limnologie* 29:123–136
- Stein ED, Brinson M, Rains MC, et al. (2009) Wetland Assessment Debate: Wetland Assessment Alphabet Soup: how to choose (or not choose) the right Assessment Method. *Soc Wetland Scientists Bull* 26:20–24
- Stelk MJ, Christie J (2014) Ecosystem service valuation for wetland restoration: what it is, how to do it, and best practice recommendations. Association of State Wetland Managers, Windham, Maine
- USDA (U. S. Department of Agriculture) (2012) Riparian Assessment: using the NRCS Riparian Assessment Method. Natural Resources Conservation Service, Bozeman, MT
- USEPA O (2020) National Ecosystem Services Classification System (NESCO) Plus. In: US EPA. <https://www.epa.gov/eco-research/national-ecosystem-services-classification-system-nescs-plus>. Accessed 10 Jun 2021
- USFWS (U.S. Fish and Wildlife Service) (2020) National Wetlands Inventory. In: National Wetlands Inventory. <https://www.fws.gov/wetlands/>. Accessed 16 Jul 2020
- Verdonschot PF (2000) Integrated ecological assessment methods as a basis for sustainable catchment management. *Hydrobiologia* 422:389–412
- Wardrop DH, Fennessy MS, Moon J, Britson A (2016) Effects of Human Activity on the Processing of Nitrogen in Riparian wetlands: implications for Watershed Water Quality. In: Vymazal J (ed) Natural and constructed wetlands: nutrients, heavy metals and energy cycling, and flow. Springer International Publishing, Cham, pp 1–22
- Wei T, Simko V (2017) R package corrplot: visualization of a correlation matrix (Version 0.84).
- Wellemeier JC, Perkin JS, Fore JD, Boyd C (2018) Comparing assembly processes for multimetric indices of biotic integrity. *Ecol Ind* 89:590–609. <https://doi.org/10.1016/j.ecolind.2018.02.024>
- Whigham DE, Lee LC, Brinson MM, et al. (1999) Hydrogeomorphic (HGM) assessment - a test of user consistency. *Wetlands* 19:560–569

Wickham H, Averick M, Bryan J, et al. (2019) Welcome to the Tidyverse. *J Open Source Softw* 4:1686. <https://doi.org/10.21105/joss.01686>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.