VALIDATION OF A RAPID WETLAND ASSESSMENT PROTOCOL FOR UTAH: EVALUATION OF SURVEY METHODS AND TEMPORAL AND OBSERVER VARIABILITY IN VEGETATION DATA

by Miles McCoy-Sulentic and Diane Menuz

REPORT OF INVESTIGATION 277
UTAH GEOLOGICAL SURVEY
a division of
UTAH DEPARTMENT OF NATURAL RESOURCES
2019

Blank pages are intentional for printing purposes

VALIDATION OF A RAPID WETLAND ASSESSMENT PROTOCOL FOR UTAH: EVALUATION OF SURVEY METHODS AND TEMPORAL AND OBSERVER VARIABILITY IN VEGETATION DATA

by
Miles McCoy-Sulentic and Diane Menuz

Cover photo: Impounded marsh in Farmington Bay Waterfowl Management Area where one of the vegetation surveys occurred.

REPORT OF INVESTIGATION 277
UTAH GEOLOGICAL SURVEY
a division of
UTAH DEPARTMENT OF NATURAL RESOURCES
2019

STATE OF UTAH
Gary R. Herbert, Governor

DEPARTMENT OF NATURAL RESOURCES
Michael Styler, Executive Director

UTAH GEOLOGICAL SURVEY
R. William Keach II, Director

PUBLICATIONS
contact
Natural Resources Map & Bookstore
1594 W. North Temple
Salt Lake City, UT 84116
telephone: 801-537-3320
toll-free: 1-888-UTAH MAP
website: utahmapstore.com
email: geostore@utah.gov

UTAH GEOLOGICAL SURVEY
contact
1594 W. North Temple, Suite 3110
Salt Lake City, UT 84116
telephone: 801-537-3300
website: https://geology.utah.gov

CONTENTS

FIGURES

TABLES

VALIDATION OF A RAPID WETLAND ASSESSMENT PROTOCOL FOR UTAH: EVALUATION OF SURVEY METHODS AND TEMPORAL AND OBSERVER VARIABILITY IN VEGETATION DATA

by
Miles McCoy-Sulentic and Diane Menuz

ABSTRACT

Plant community composition data are often an important component of ecological studies, but their value is only as good as the appropriateness of the methods used to collect the data and the metrics used to summarize the data. Testing the degree of tradeoffs among methods and the degree of sensitivity of metrics is a valuable exercise to help select methods and metrics that are most appropriate for project goals. In 2017, the Utah Geological Survey (UGS) undertook an effort to compare different vegetation survey methods for use in its Utah Rapid Assessment Protocol (URAP), a rapid assessment method used to assess the condition and potential function of wetlands in Utah. Though vegetation data collected as part of URAP have been used to calibrate the URAP condition assessment, no analysis has been conducted to compare other possible survey methods or to evaluate its repeatability among surveyors or during different survey periods.

We evaluated three different methods for collecting URAP vegetation data—a progressive timed meander, four 100-m^2 plots, and line point-intercept (LPI)—and examined variability among methods and between observers. Differences in species detection, cover estimates, and 10 common vegetation metrics based on absolute cover, relative cover, and mean C, were used to assess variability. Inter-observer variability was not tested in LPI due to the greater time requirement. Each of the three methods and the repeated visits were conducted at six sites, and the meander was conducted at 19 sites. Surveys from previous years were used to compare the 100-m^2 plot method to the meander and to compare meander results across years.

Inter-observer differences in cover estimates and species lists were smaller in the 100-m^2 plots than the meander survey. Pseudoturnover, a measure of differences between observation efforts, was a source of differences within and across years though rates between observers were similar to other studies. Data from the meander survey and 100-m^2 plots were strongly correlated, and both captured more species than LPI. The meander and 100-m^2 plots both recorded significantly lower absolute cover estimates than LPI. Vegetation metrics calculated from relative cover estimates and mean C showed smaller differences than metrics based on absolute cover and mean weighted C and resulted in similar site rankings across all methods. Time of visit did not have an effect on vegetation metrics while survey methods resulted in significant differences in absolute cover and richness metrics.

We recommend the use of the meander method for URAP vegetation surveys because it captures greater species diversity, is comparable to more intensive methods, and requires less time and cost. Relative cover-based metrics and mean C are more repeatable than absolute cover and weighted C metrics and will be used in future analyses. Other methods should be employed if absolute cover metrics are to be used. Differences caused by species misidentification may be mitigated through increased training of staff and encouraging a greater number of field collections.

INTRODUCTION

Plant community composition data are often an important component of ecological studies, but their value is only as good as the appropriateness of the methods used to collect the data and the metrics used to summarize the data. There can often be tradeoffs between the accuracy of a given method of collecting vegetation data and the time needed to collect the data. Methods that are more accurate for obtaining plant cover estimates often use small plots or individual points as the basic sampling unit but can require a large number of units to obtain estimates across an entire site. In contrast, methods that use larger plots may be faster and more likely to capture uncommon species since they survey larger areas but can have lower accuracy and higher variability among surveyors. Plant community composition data are typically summarized into metrics that describe ecological components of interest such as species richness or cover of noxious weed species. Some metrics may be more sensitive to changes within a growing season (e.g., total cover, cover of a late-growing species) while other metrics may be more stable despite observer and seasonal variability. For rapid assessments, an ideal survey method is one that may be quickly implemented, adequately capture species diversity, and may be calibrated to more intensive survey methods. An ideal vegetation metric is one that is relevant to the study and is relatively stable across observers and within a season.

The Utah Geological Survey (UGS) developed the Utah Rapid Assessment Procedure (URAP) in 2014 as a tool to rapidly assess the condition of Utah's wetland resources. URAP is intended to provide basic inventory information on the condition and potential function of Utah's wetlands and has been implemented in several regions in the state (Menuz, 2016a, 2016b; Menuz and others, 2016; Menuz and Sempler, 2018). Vegetation data are collected for URAP using a meander method, in which an observer walks around an approximately 0.5 ha plot, records all observed species, and then estimates the cover of each recorded species. In theory, the survey should be completed in one hour, but observers often spend more time on the survey because of the large number of plant species encountered or because wetlands are difficult to walk around (due to deep water, dense vegetation, etc.). Vegetation data are used to calculate metrics that help evaluate wetland condition, and to help develop wetland reference standards. The method assumes that the person collecting vegetation data is a well-trained botanist familiar with wetland plants in the western United States.

Though URAP vegetation data have been tested against other measures of wetland condition (Menuz and others, 2016), no analysis has been conducted to compare the meander method to other possible methods or to evaluate its repeatability among surveyors or during different times in the growing season. While differences between observers are to be expected in vegetation surveys even among experienced botanists (Sykes and others, 1983; Lepš and Hadincová, 1992; Vittoz and Guisan, 2007), it is valuable to quantify those differences to guide method selection and evaluate the appropriateness of vegetation metrics used in analysis. Furthermore, metrics calculated from URAP vegetation data have not been evaluated to determine their sensitivity to variability between observers or to differences within a growing season (for sites visited only once in a growing season).

The UGS undertook a field study in 2017 to evaluate different methods for collecting URAP vegetation data to quantify variability among methods and between observers in species lists, species cover estimates, and in vegetation metrics commonly used in condition assessments. The goal of this study was to identify a suitable survey method that was both rapid and consistent between observers, to identify vegetation metrics that were relatively stable across observers and times of the year, and to better understand possible errors and discrepancies between surveyors. We compared results of three methods: a progressive timed meander, four 100-m^2 plots (in an arrangement similar to the U.S. Environmental Protections Agency's (EPA) National Wetland Condition Assessment (NWCA) method (USEPA, 2016; hereafter referred to as NWCA plots), and line point-intercept (LPI) and evaluated 10 vegetation metrics. We also examined inter-observer variability in the timed meander and NWCA plots and the temporal variability of the timed meander method across and within years, hereafter referred to as inter- and intra-annual variability, respectively.

METHODS

Survey Sites

We collected vegetation data at 19 sites in 2017, all of which had been surveyed by the UGS as part of watershed surveys in 2014-2016 (Menuz and others, 2016; Menuz and Sempler, 2018). We selected the 19 sites from the list of previously surveyed sites with the aim of having equal representation of four wetland types, including alkaline depressions, marsh, wet meadows, and montane shrublands. Within these classes, we selected sites that were at least 0.3 ha in size and represented a range of vegetation conditions, requiring that sites have a minimum of five species and represented low, medium, and high diversity for the wetland type. We gave preference to sites that were previously surveyed in June or July to align with our 2017 survey period, were within 50 miles of the office (as the crow flies) and within ½ mile of the nearest road for ease of access, and had NWCA plot data previously recorded. We surveyed five wetlands in each wetland class except for montane shrublands, where only four sites were sampled due to time constraints. We collected timed meander data at all 19 sites to evaluate inter-observer and inter-annual variability and used six of the 19 sites to evaluate survey methods and intra-annual variability (figure 1). The six sites used for method testing included two meadows, two alkaline depressions, one marsh, and one montane shrubland and were selected because they were all standard 40-m circular assessment plots. Intra-annual testing of the meander method was conducted at both circular and freeform sites at two meadow, two shrubland, one alkaline depression, and one marsh site, which included one of the six sites used for the method testing. For sites visited more than once, we used data from the first site visit in the analysis of inter-observer and inter-annual variability. We used data from 21 sites surveyed in 2014 (Menuz and others, 2016) and 2017 in an additional comparison of the meander and NWCA plot methods. All sites were located in northern Utah near Salt Lake City, in either the Central Basin and Range or Wasatch and Uinta Mountains ecoregions (Omernik, 1987).

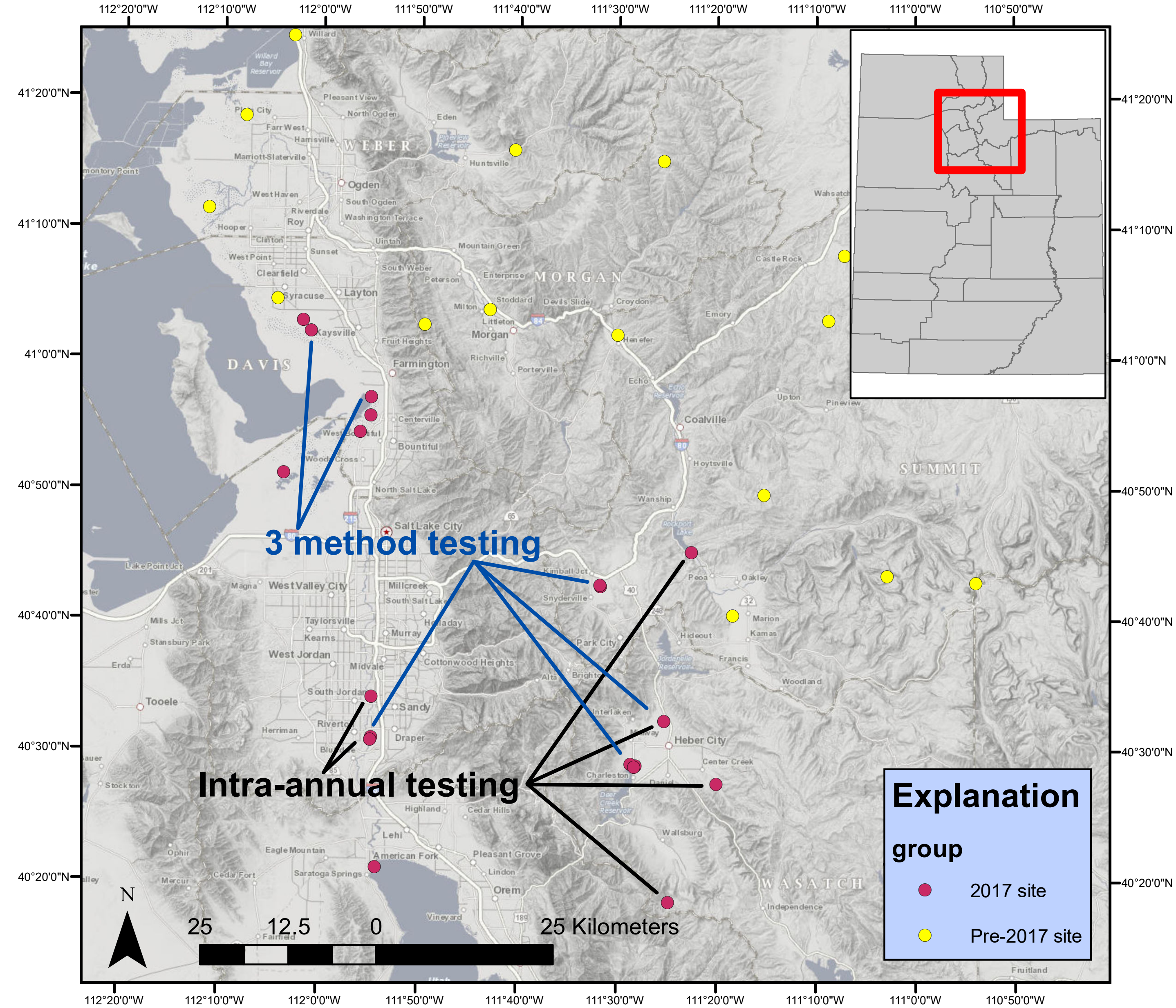

Figure 1. *Sites surveyed for the wetland assessment vegetation protocol testing, labeled by the survey type and year of survey. All sites were surveyed using a meander method.*

Field Methods

For this project, surveyors set up each site as one of three plot types to conduct a timed meander survey: 40-m circular, rectangular, or freeform, based on the plot type used in the site's previous survey. Circular plots were established by running a transect tape 40 m in each of the four cardinal directions from the plot center point to flag the boundary, and rectangular plots were established by flagging the corners and midpoints of the sides. Freeform plots were established using spatial data from a GPS unit to determine and flag site boundaries. All plots were between 0.38 and 0.64 ha in area. Flagging was removed between survey visits at the sites surveyed multiple times during the growing season.

The meander method was adapted from the Minnesota Pollution Control Agency's Rapid Floristic Quality Assessment (Minnesota Pollution Control Agency, 2014) and differed from the method used in previous URAP surveys by requiring surveyors to continue surveying if they found many new species during the last 10 minutes of the allotted survey time. Surveyors started with a base survey time of 30 minutes for the first plant community present at a site plus 20 additional minutes for each additional community present. Communities were identified as distinct groupings of species with similar physiognomy (e.g., wet meadow, shrub complex). Next, surveyors walked around each plant community and recorded every plant species they encountered. Surveyors "meandered" around the plot and could focus their effort on visiting a variety of microhabitats within the plot rather than being confined to transects or plots. If three or more species were found in the last 10 minutes of the base survey, an additional 10 minutes were added. Additional 10-minute increments were added as needed until less than three new species were encountered in the final 10 minutes. Surveyors estimated the absolute cover of each species at the end of the survey. To test inter-observer variability, two surveyors independently conducted each meander survey without consulting with one another about plant identification or cover estimates. After recording their cover estimates, surveyors discussed their results with one another and came up with a consensus list of plant species and cover estimates for the entire site. The consensus data were used for the inter- and intra-annual variability and method testing rather than data from either of the individual surveyors.

At six of the 40-m circular plot sites, we also collected vegetation data using NWCA plots and LPI transects (figure 2). For the NWCA plot method, we established four 100-m^2 plots in a manner used by the UGS in previous watershed surveys (Menuz and others, 2016) and similar to the arrangement described by the EPA's NWCA protocol (USEPA, 2016). Surveyors set up 10 m x 10 m plots 20 m to the east, 15 m to the north, 15 m to the west, and 5 m to the south of the site center, and flagged the four corners of each NWCA plot to aid with the assessment. The two surveyors worked together to set up each NWCA plot, but recorded species and cover estimates independently. Surveyors then came up with a consensus list of species and cover values once they both surveyed all four plots.

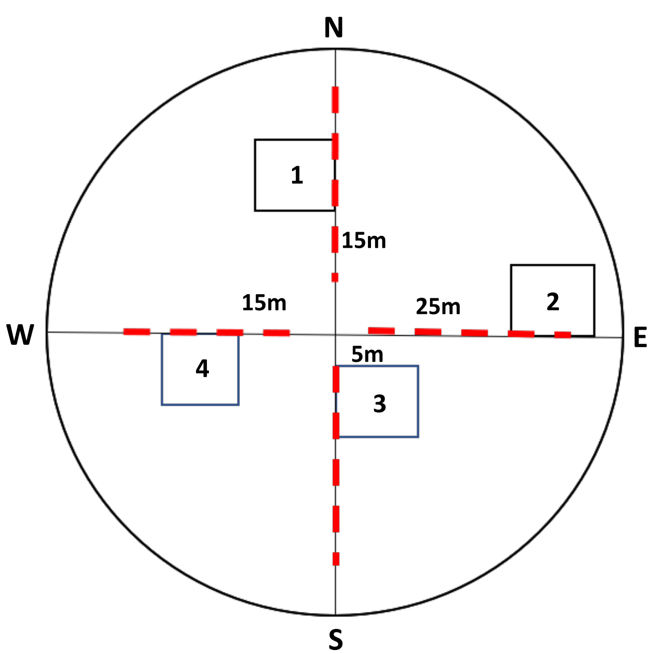

Figure 2. *Plot setup for meander, NWCA plots, and line point-intercept for circular assessment plot. Circle represents 0.5 ha plot for meander, numbered boxes represent 100-m^2 NWCA plots, and dashed lines at cardinal directions indicate line point-intercept transects.*

Line point-intercept (LPI) was conducted along the four transects used to establish the NWCA plots in the four cardinal directions. For this method, we dropped a pinflag along the transect at 0.5 m intervals beginning 5 m from the center of the plot to avoid trampling caused during plot center establishment and extending for 25 m for a total of 50 points per transect and 200 points per site. Each plant species touching the pin was recorded at each point. Both observers worked together to conduct the LPI survey, with one observer and one recorder. Differences between observers were not examined for LPI due to the more time-intensive requirements of the method and because the method has been found in other studies to be more objective and repeatable (Everson and Clarke, 1987; Vittoz and Guisan, 2007).

Species not able to be identified in the field were collected and identified in the office. Final determinations of collected specimens from both observers were made after all fieldwork was completed and species lists at each site were updated and resolved before analysis. Thus, the intra-annual analysis does not analyze discrepancies in species lists due to differences in identification of the same species between observers, only differences in species detection and cover estimation. However, as different botanists conducted fieldwork during different years, comparison of inter-annual data does include discrepancies in species lists due to differences in identification. All field forms are available in appendix A.

Data Analysis

We converted NWCA plot and LPI vegetation data to overall site cover and richness estimates before analysis and used meander data as they were originally recorded. NWCA species cover estimates were converted by taking the mean of all four plots. LPI cover estimates were calculated by dividing the total number of times a species was recorded at a site by the total number of points (200) to calculate a percent cover estimate for each species. Consensus estimates from the two surveyors were used in the inter-annual, intra-annual, and between methods analysis. All calculations and statistical tests were done using R statistical software (R Core Team, 2018).

General differences in the vegetation surveys between observers and methods (i.e., observer 1 to observer 2, visit 1 to visit 2, method 1 to method 2) were explored using basic summary statistics and by calculating a mean pseudoturnover rate. Pseudoturnover rate is a measure of difference in species detection and was calculated following the formula described by Nilsson and Nilsson (1985) which is defined as:

$$PT = ((A+B)/(S_A+S_B)) *100$$

where A and B represent the numbers of species found by only one observer or one method or on only one visit, and S_A and S_B are the total number of species found by the corresponding observer (or method or visit).

We used vegetation data to calculate 10 metrics that summarize important attributes of wetland plant community composition (table 1), and explored the variability in these metrics between observers, methods, within a season, and across years. These metrics were chosen because they were metrics of interest to partner agencies, were strongly correlated with wetland condition in other studies (Menuz and others, 2016), or are used for scoring condition metrics in the URAP protocol. These metrics summarize species richness of different functional groups (i.e., herbs, shrubs, trees), and plant community characteristics such as hydrologic affinity, nativity, and disturbance. Most metrics rely on species cover, though three of the metrics are related to species richness including shrub richness, herbaceous (non-woody) richness, and overall species richness; tree species were largely absent from sites and thus tree richness was not evaluated. Metrics involving cover by wetland species included only species with wetland indicator status ratings of OBL and FACW. Comparisons of shrub richness were largely excluded from intra-annual and method comparisons as most sites used for these comparisons did not have shrubs.

Two of the metrics rely on plant coefficient of conservatism values, or C-values, for their calculation. C-values between 1 and 10 are assigned to species based on their association with disturbance through a combination of best professional judgment, literature review, and field observations. Low values indicate that species are usually found at disturbed sites, high values indicate that species are associated with pristine sites, and values in the middle indicate that species may be found equally at either type of site. All non-native species are assigned a C-value of 0. C-values are typically developed for individual states or regions to capture regional variability in how species respond to disturbance, but C-values have not been developed for the state of Utah. We instead assigned species the mean C-value from four surrounding states; Colorado (Rocchio, 2007), Montana (Jones, 2005), Wyoming (Washkoviak and others, 2015), and Idaho, which uses C-values from eastern Washington's Columbia Basin region (Rocchio and Crawford, 2013). To adjust for use in Utah we made sure that every non-native species, and no native species, had a C-value of 0. This list was last updated in 2017 (Menuz and Sempler, 2018).

We calculated the mean and standard deviation of the absolute values of differences in metric values for all comparisons (inter-observer, inter-annual, intra-annual, and between methods), as well as Spearman rank correlation to determine if site rankings were preserved. Paired t-tests were used to determine if inter-observer differences were greater in the NWCA or meander method at six sites, and if consensus values differed significantly between NWCA and meander methods at 21 inter-annual

Table 1. *Vegetation metrics evaluated in this study, where x_i = percent cover for i^{th} species, C_i = coefficient of conservatism, W_i = percent cover of OBL and FACW species, N_a = percent cover of native species, N_x = percent cover of noxious weeds, N_t = count of all species, N_h = count of herbaceous species, N_s = count of all shrubs.*

Metric	Description	Calculation
Absolute wetland cover	Total cover of only OBL and FACW species	$\sum_{i=1}^{n} W_i$
Absolute native cover	Total cover of only native species	$\sum_{i=1}^{n} N_a$
Absolute noxious cover	Total cover of only noxious weed species	$\sum_{i=1}^{n} N_x$
Relative wetland cover	Cover of OBL and FACW species as a percent of total vegetation cover[1]	$\sum_{i=1}^{n} W_i / \sum_{i=1}^{n} x_i$
Relative native cover	Cover of Native species as a percent of total vegetation cover[1]	$\sum_{i=1}^{n} N_a / \sum_{i=1}^{n} x_i$
Mean C	Mean C-value of all species[1]	$\sum_{i=1}^{n} (C_i) / N_t$
Mean weighted C	Cover-weighted mean C-value of all species[1]	$\sum_{i=1}^{n} (x_i C_i) / \sum_{i=1}^{n} x_i$
Shrub richness	Total number of shrubs	N_s
Species richness	Total number of all species	N_t
Herbaceous richness	Total number of herbaceous species (all non-woody species)	N_h

[1]Calculated using only species with known wetland indicator, nativity status, or C-value.

sites. To test if time of visit or survey method had a significant overall effect on metric values, we used linear mixed effects models with visit number (or method) as a fixed effect and site as a random effect using the *lmerTest* package (Kuznetsova and others, 2017) with pairwise contrasts calculated using the *emmeans* package (Lenth, 2019). Metrics not able to meet model assumptions of equal variance and normality were log transformed for analysis though data are presented in their original units of measurement.

We excluded some site and metric combinations with missing data from the inter-annual comparison. We set a requirement that at least 80% of a site's plant species have an assigned C-value for the site's mean C metric to be included in analysis. Similarly, we required that 80% of a site's plant cover have an assigned C-value, nativity, or wetland indicator status to be included in the analysis of the corresponding metrics. This approach was taken for the inter-annual data to remove unidentified species and make the data more comparable, but not for the intra-annual and inter-observer comparison as unidentified species were the same for both observers. We also analyzed how variability in two vegetation metrics led to variability in corresponding URAP condition metrics, which is presented in appendix B.

RESULTS

Inter-Observer Variability

Meander Survey

The mean pseudoturnover rate between two observers at 19 sites using the meander method was 19.6% and ranged from 7.7% to 33.3%. Both observers recorded the same species 67% of the time. Ten of the 165 species recorded by only one observer had >1% cover, and none had >7% cover. Differences in absolute cover estimates between two observers ranged from 0 to 40% cover, though estimates were within 13% of one another 95% of the time. Observers were in perfect agreement on cover estimates 28% of the time.

Absolute cover metrics had larger mean differences in estimates between observers than their relative cover counterparts (table 2). For example, absolute and relative wetland cover had mean differences between observers of 15.28% and 7.88%, respectively. Mean difference between observers was slightly greater for mean weighted C than mean C (0.38 and 0.18, respectively). Species richness and herbaceous richness were similar with mean differences of 2.89 and 2.58

Table 2. *Mean inter-observer differences with standard deviation (SD) and Spearman rank correlation coefficient (r_s) for meander and NWCA plot surveys. Paired t-tests performed on differences between surveyors at the six sites with both NWCA plot and meander data.*

Metric	Meander (n=19)		NWCA Plot (n=6)		Meander (n=6)		Paired t-test p-value
	Mean difference (SD)	r_s (p-value)	Mean difference (SD)	r_s (p-value)	Mean difference (SD)	r_s (p-value)	
Absolute wetland cover (%)	15.28 (12.07)	0.68 (<0.01)	10.67 (12.14)	0.60 (0.24)	15.23 (10.55)	0.83 (0.06)	0.21
Absolute native cover (%)	14.17 (11.44)	0.82 (<0.01)	9.28 (7.85)	0.83 (0.06)	14.62 (12.21)	0.71 (0.14)	0.37
Absolute noxious cover (%)	5.48 (7.11)	0.93 (<0.01)	1.80 (2.68)	0.81 (0.05)	4.70 (4.53)	0.77 (0.10)	0.24
Relative wetland cover (%)	7.88 (5.72)	0.94 (<0.01)	5.39 (4.1)	1.00 (0.02)	11.7 (5.87)	0.94 (0.02)	0.05
Relative native cover (%)	8.98 (6.75)	0.89 (<0.01)	5.69 (3.57)	0.94 (0.02)	11.1 (6.81)	0.49 (0.36)	0.24
Mean C	0.18 (0.13)	0.93 (<0.01)	0.16 (0.09)	0.94 (0.02)	0.16 (0.05)	1.00 (<0.01)	0.94
Mean weighted C	0.38 (0.34)	0.86 (<0.01)	0.19 (0.15)	0.94 (0.02)	0.54 (0.45)	0.31 (0.56)	0.17
Shrub richness	0.21 (0.54)	1.00 (<0.01)	[1]Not performed				
Species richness	2.89 (1.7)	0.93 (<0.01)	2.00 (1.26)	0.94 (<0.01)	2.67 (1.21)	0.84 (0.04)	0.33
Herbaceous richness	2.58 (1.39)	0.92 (<0.01)	2.17 (0.98)	0.99 (<0.01)	2.67 (1.21)	0.87 (0.02)	0.42

[1]Not performed due to low number of sites with shrubs.

species, respectively. Shrub richness showed a mean difference of 0.21, though most sites contained no shrubs. Observers' metric values were generally strongly correlated, with Spearman rank correlation (r_s) for all metrics $r_s \geq 0.82$ except for absolute wetland cover (r_s=0.68).

NWCA Plots

The mean pseudoturnover rate between observers using the NWCA plots at six sites was 11.4%, lower than the rate between two observers using the meander method at both the same six sites (21.3%) and all 19 sites (19.6%). Overall, species were recorded by both observers 80% of the time in the plots versus 65% of the time in the meander survey at the same six sites. Most of the observations made by only one observer were for species with ≤1% cover, with only two cases having greater cover, though still <3%. Differences in absolute cover estimates for individual species ranged from 0 to 21.3%, though estimates were within 6.3% of one another 95% of the time. Observers were in perfect agreement on cover estimates 13% of the time.

Observer differences in metrics for NWCA data were similar to patterns in the meander survey, with smaller mean differences in relative cover metrics than absolute cover metrics, and mean differences in herbaceous and species richness less than three (table 2). Observer estimates were strongly correlated ($r_s \geq 0.81$), with the exception of absolute wetland cover (r_s=0.60). Observers differed more in their estimates of relative wetland cover using meander surveyed data compared to NWCA plot data based on a paired t-test; other differences were not significant.

Variability Between Visits

Intra-Annual

The mean pseudoturnover rate for intra-annual meander surveys was 17.5%. Overall, 40% of species were recorded during all three visits, 24% during two visits, and 35% during only one visit. While 88% of species that were only recorded once or twice had <1% cover, 12 of these observations had between 1% and 7% cover and four had between 17% and 30% cover. These latter four species were all grasses that may have been confused with other grass species during some of the site visits. For example, *Dechamspia caespitosa* was assumed to be the dominant grass at a site early in the season before most grasses were flowering, but later surveys showed that other grass species were more common. The very similar grasses *Agrostis stolonifera* and *Agrostis gigantea* were also recorded once each at the same site with 25% and 28% cover, respectively, suggesting likely misidentification on one of the visits. More than half (11 of 18) of the largest discrepancies in cover (differences > 5%) were for graminoid species (grasses, rushes, sedges), with most of those involving grasses. Though some variability in metrics was seen at some sites (figure 3), time of visit did not have a significant overall effect on vegetation metric values during three repeated visits during the growing season based on linear mixed effects models (table 3). Differences in metrics between visits showed a similar pattern to the inter-observer differences, with relative cover metrics showing lower variability than absolute cover metrics, relatively low mean differences in richness values, and differences in mean C ≤0.31 (table 4). Vegetation metric values at a site were strongly correlated across all visits (r_s>0.83) with the exception of absolute native cover between the first and second visits (r_s=0.66), which was explained by species misidentification of *Dechamspia caespitosa* described above.

Inter-Annual

Mean pseudoturnover between survey years was 46.2%, much greater than seen in other comparisons. While surveyors recorded the same species across years only 38% of the time, most species recorded only once had <1% cover; 18% of these cases had cover values between 2% and 60%. Inter-annual comparisons may include differences in species lists due to differences in detection as well as differences in identification from different botanists, and true turnover or population change that may have occurred in the course of up to three years. Many of the observations with the largest differences in cover estimates were likely due to differences in species identification of related species (e.g., *Carex aquatilis* versus *Carex nebrascensis*) or where one team identified a species to genus only. For species recorded during both years, 59% of cover estimates differed by ≤1%, 32% differed between 1% and 10%, and 9% had differences >10%. The largest differences in cover estimates between years were observed for *Lemna minor* and *Phragmites australis* and likely reflected true changes in abundance rather than differences in observer estimation. For example, *Lemna minor* (a floating, obligate wetland species) was recorded in a flooded impoundment with 60% cover in 2015 and not recorded two years later when the impoundment was drained and completely dry. At another site, estimated cover of the invasive grass *Phragmites australis* increased 62% in two years. Photos taken during site visits confirm this dramatic change and provide evidence of managed burning to control invasive *Phragmites australis* prior to the first visit.

As seen in other comparisons, metrics involving relative cover showed smaller differences between surveys than metrics involving absolute cover. For example, mean difference in relative wetland cover was more than three times less than absolute wetland cover (7.40% vs 27.55%, table 4). Mean differences in absolute wetland cover, absolute native cover, and absolute noxious cover were larger across years than both between observers and at different times within a year. Relative and mean C metrics showed similar differences in inter-annual, intra-annual, and inter-observer comparisons, while

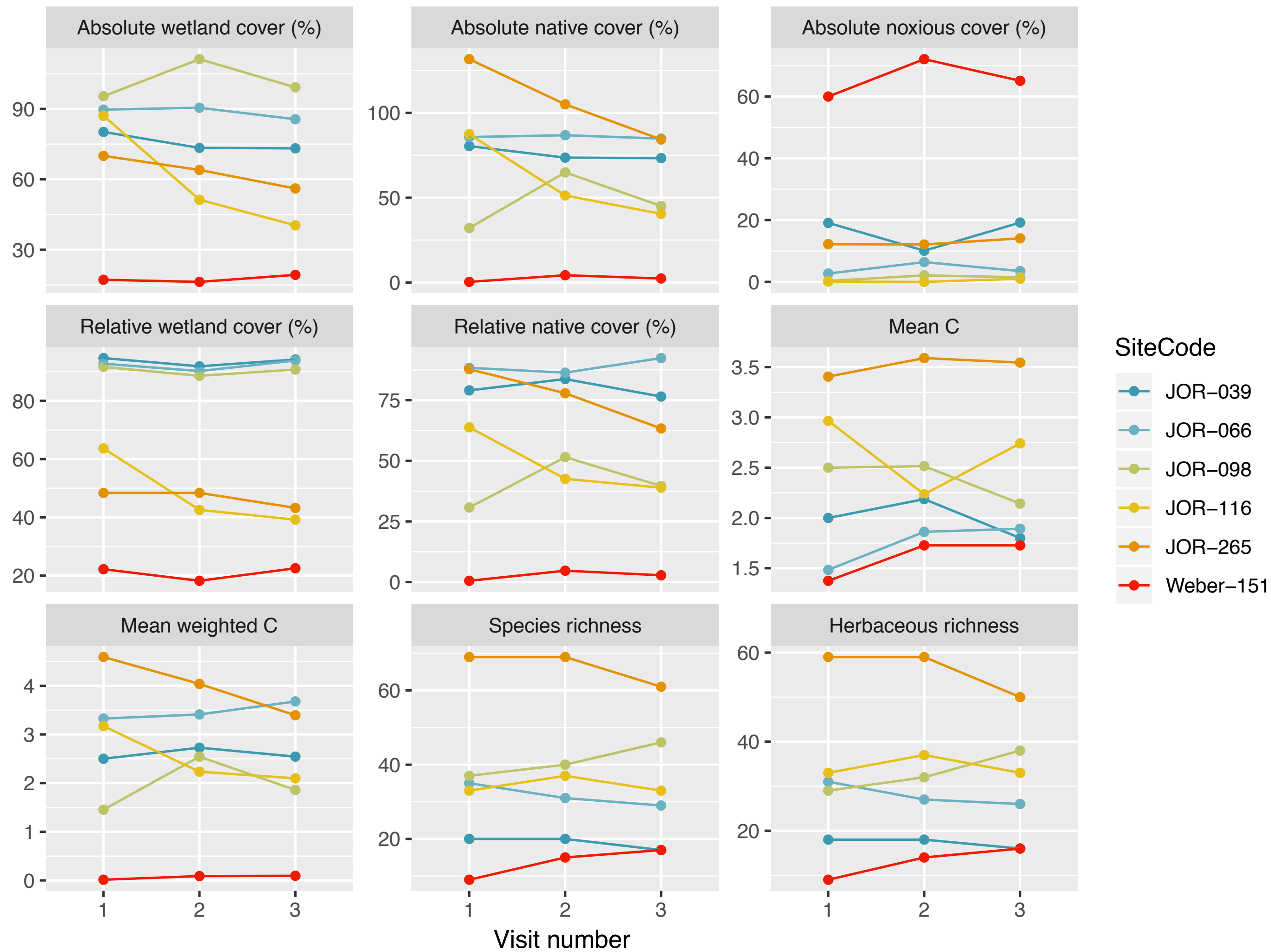

Figure 3. *Vegetation metric values at each of three intra-annual visits by visit number and site, calculated with consensus data. Time of visit did not have an overall significant effect on any metric examined based on linear mixed effects model.*

Table 3. *Model structure and relevant output of linear mixed effects models for intra-annual and 3 method testing.*

Intra-annual linear mixed effects model: Metric ~ Visit # + (1|site)				
Metric	**Variable**	**Estimate**	**Std. error**	**P value**
Absolute wetland cover	Intercept (Visit 1)	33.83	7.72	<0.01
	Visit 2	1.50	2.20	0.51
	Visit 3	~ 0	2.20	1.00
Absolute native cover	Intercept (Visit 1)	69.60	15.60	<0.01
	Visit 2	-5.28	8.70	0.56
	Visit 3	-14.53	8.70	0.13
Absolute noxious cover	Intercept (Visit 1)	1.89	0.60	0.02
	Visit 2	0.17	0.18	0.36
	Visit 3	0.28	0.18	0.15
Relative wetland cover	Intercept (Visit 1)	69.89	12.71	<0.01
	Visit 2	-5.58	3.08	0.10
	Visit 3	-4.92	3.08	0.14

Table 3. *Continued.*

Relative native cover	Intercept (Visit 1)	59.38	13.51	<0.01
	Visit 2	-0.59	5.16	0.91
	Visit 3	-6.12	5.16	0.26
Mean C	Intercept (Visit1)	3.29	0.30	<0.01
	Visit 2	0.06	0.15	0.67
	Visit 3	0.02	0.15	0.89
Mean weighted C	Intercept (Visit 1)	3.51	0.58	<0.01
	Visit 2	~ 0	0.25	0.99
	Visit 3	-0.23	0.25	0.37
Species richness	Intercept (Visit 1)	34.83	7.17	<0.01
	Visit 2	1.50	2.20	0.51
	Visit 3	<0.01	2.20	1.00
Herbaceous richness	Intercept (Visit 1)	30.83	6.34	<0.01
	Visit 2	1.33	2.18	0.55
	Visit 3	~ 0	2.18	1.00
3 Method testing linear mixed effects model: Metric ~ Method + (1|site)				
Metric	**variable**	**Estimate**	**Std. error**	**P value**
Absolute wetland cover	(intercept) LPI	101.75	13.02	<0.01
	Meander	-32.35	5.56	<0.01
	NWCA	-47.30	5.56	<0.01
Absolute native cover	(intercept) LPI	102.25	8.01	<0.01
	Meander	-33.00	6.64	<0.01
	NWCA	-45.83	6.64	<0.01
Absolute noxious cover	(intercept) LPI	1.75	0.58	0.02
	Meander	0.26	0.31	0.42
	NWCA	-0.16	0.31	0.61
Relative wetland cover	(intercept) LPI	69.84	11.43	<0.01
	Meander	-5.09	4.51	0.29
	NWCA	-2.78	-0.62	0.55
Relative native cover	(intercept) LPI	70.87	7.88	<0.01
	Meander	-7.32	5.30	0.20
	NWCA	-4.73	5.30	0.39
Mean C	(intercept) LPI	2.34	0.38	<0.01
	Meander	0.04	0.16	0.82
	NWCA	0.09	0.16	0.56
Mean weighted C	(intercept) LPI	2.74	0.31	<0.01
	Meander	-0.27	0.24	0.29
	NWCA	-0.10	0.24	0.69
Species richness	(intercept) LPI	13.67	2.75	<0.01
	Meander	13.67	1.87	<0.01
	NWCA	6.67	1.87	<0.01
Herbaceous richness	(intercept) LPI	12.50	2.42	<0.01
	Meander	13.17	1.71	<0.01
	NWCA	6.50	1.71	<0.01

Table 4. *Mean differences with standard deviation (SD) between visits for intra-annual and inter-annual comparisons using consensus data, with Spearman rank correlation (r_s) and p-value. Six sites were used for intra-annual comparison, number of sites used for inter-annual comparison indicated by* n.

Metric	Intra-annual						Inter-annual		
	Visit 1 v 2		Visit 2 v 3		Visit 1 v 3		Year 1 v 2		
	Mean difference (SD)	r_s (p-value)	Mean difference (SD)	r_s (p-value)	Mean difference (SD)	r_s (p-value)	*n*	Mean difference (SD)	r_s (p-value)
Absolute wetland cover (%)	11.02 (13.32)	0.83 (0.06)	6.48 (4.61)	1.00 (<0.01)	12.93 (17.06)	1.00 (<0.01)	15	27.55 *(21.47)*	0.62 (0.02)
Absolute native cover (%)	17.88 (15.69)	0.66 (0.18)	9.25 (9.29)	0.94 (0.02)	19.53 (21.78)	0.94 (0.02)	18	25.96 *(20.55)*	0.47 (0.05)
Absolute noxious cover (%)	4.47 (4.99)	0.94 (0.02)	3.77 (3.48)	0.94 (0.02)	1.67 (1.78)	0.94 (0.02)	19	12.47 *(18.00)*	0.63 (<0.01)
Relative wetland cover (%)	5.58 (7.72)	0.94 (0.02)	3.49 (1.15)	1.00 (<0.01)	5.40 (9.50)	1.00 (<0.01)	15	7.40 *(9.55)*	0.90 (<0.01)
Relative native cover (%)	10.46 (8.58)	0.89 (0.03)	7.53 (4.85)	1.00 (<0.01)	11.18 (10.72)	1.00 (<0.01)	18	10.61 *(12.02)*	0.74 (<0.01)
Mean C	0.31 (0.25)	0.94 (0.02)	0.22 (0.22)	0.89 (0.03)	0.28 (0.11)	0.89 (0.03)	13	0.22 *(0.18)*	0.86 (<0.01)
Mean weighted C	0.49 (0.44)	0.83 (0.06)	0.32 (0.28)	0.89 (0.03)	0.53 (0.50)	0.89 (0.03)	17	0.57 *(0.35)*	0.65 (<0.01)
Shrub richness	0.5 (0.84)	0.94 (<0.01)	0.5 (0.55)	0.91 (0.01)	0.67 (0.52)	0.91 (0.01)	19	0.42 *(1.02)*	0.99 (<0.01)
Species richness	2.83 (2.40)	0.94 (0.02)	4.17 (2.40)	0.99 (<0.01)	5.67 (3.50)	0.99 (<0.01)	19	6.05 *(6.48)*	0.83 (<0.01)
Herbaceous richness	2.67 (2.16)	0.94 (0.02)	4.00 (3.03)	0.93 (<0.01)	5.33 (3.72)	0.93 (<0.01)	19	5.53 *(5.37)*	0.76 (<0.01)

both richness metrics were similar to at least one intra-annual comparison, though greater than inter-observer differences (tables 2 and 4). All metrics were significantly correlated between years but showed weaker correlation for absolute cover metrics and mean weighted C (all $r_s \leq 0.65$) than for relative cover metrics and mean C (all $r_s \geq 0.74$). Richness metrics were all significant and strongly correlated (all $r_s \geq 0.75$).

Variability Between Methods

The mean psuedoturnover rate was highest between meander and LPI, at 37.8%, while the meander and NWCA, and NWCA and LPI method mean psuedoturnover rates were 26.6% and 24.2%, respectively. Species were recorded by all three methods 40% of the time. Ninety-five percent of the species recorded by only one method had ≤1% cover and 95% of species recorded by only two methods had ≤3% cover.

Survey method had a significant overall effect on metric values for absolute wetland cover, absolute native cover, species richness, and herbaceous richness based on linear mixed effects models (table 3). LPI estimates for absolute wetland and absolute native cover were significantly higher than estimates using the other two methods (figure 4). All three methods produced significantly different results from one another for species and herbaceous richness, with the meander recording the greatest number of species and LPI the least for both metrics. The meander method found nearly twice the number of species as LPI at five of the six sites. Shrub richness was not analyzed due to the low number of sites with shrubs. Rank ordering of plots based on vegetation metrics was mostly maintained ($r_s \geq 0.83$) with exceptions for some comparisons of absolute native cover, mean weighted C, and herbaceous and species richness (table 4). Mean differences were much greater between methods than between observers for absolute wetland cover, absolute native cover, and both richness metrics (tables 2 and 5).

Additional comparison between NWCA and meander methods using 21 sites showed smaller mean differences for all metrics than for the set of six sites for all vegetation metrics (table 5). The 21 sites had significantly more herbaceous and total species recorded in the meander method than the plots, based on paired t-tests. The meander method also pro-

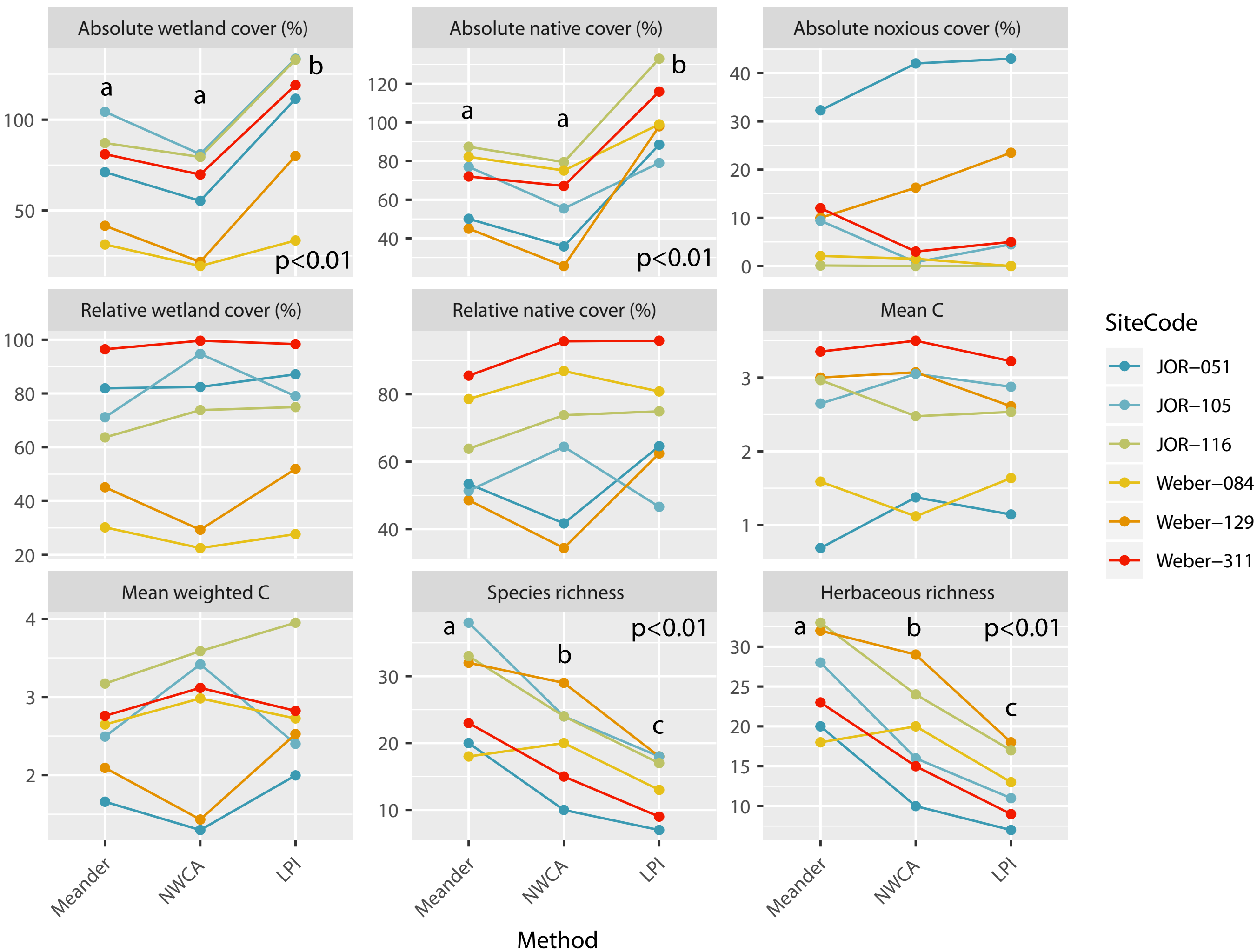

Figure 4. *Vegetation metrics at the six method testing sites based on consensus data from three different survey methods. P-values from linear mixed effects models are reported only for metrics where method had an overall significant effect (p<0.05). Letter codes indicate pair-wise comparisons with p<0.05.*

duced significantly greater cover estimates of absolute native, absolute wetland, and absolute noxious cover. Metrics obtained using the two methods were strongly correlated, with all metrics showing significant rank correlation coefficients between 0.86 and 0.95, suggesting that rank order was largely preserved between methods.

DISCUSSION

Inter-Observer Variability

Though species lists differed between observers, pseudoturnover rates were comparable to those reported from other studies (Morrison, 2016). The majority of species discrepancies in our data were due to very low cover species and did not translate to large differences in richness estimates or mean C values. This suggests that mean C is robust to differences in species lists, though some caution should be taken at very low richness sites where means would be more sensitive to inclusion or exclusion of a few species. Mean differences in relative cover metrics were smaller than mean differences in absolute cover metrics suggesting that metrics based on relative cover are more repeatable and may help account for differences in observer cover estimates.

Pseudoturnover rates and mean differences in metric values between observers were almost always lower in the NWCA plots than the meander survey, though differences were only significant for relative wetland cover (possibly due in part to low sample size). This result was unsurprising because the meander method required the observer to survey 0.5 ha whereas the NWCA method covers 0.04 ha in four smaller plots, though Vittoz and Guisan (2007) observed the opposite pattern, with pseudoturnover increasing with smaller plot size. Inter-observer differences were smaller between methods for mean C than for mean weighted C. Though inter-observer

Table 5. Mean differences with standard deviation (SD) between three survey methods tested at six and twenty-one sites, with Spearman rank correlation coefficient (r_s) and p-value. Paired t-test p-values indicate whether vegetation metric values differed significantly between methods for data from twenty-one sites only.

Metric	3 method comparisons (n=6)						Additional 2 method comparison (n=21)		
	LPI v NWCA		LPI v Meander		NWCA v Meander		NWCA v Meander		
	Mean difference (SD)	r_s (p-value)	Mean difference (SD)	r_s (p-value)	Mean difference (SD)	r_s (p-value)	Mean difference (SD)	r_s (p-value)	Paired t-test p-value
Absolute wetland cover (%)	47.30 (16.58)	1.00 (<0.01)	32.35 (15.72)	1.00 (<0.01)	14.95 (5.85)	1.00 (<0.01)	9.80 (8.10)	0.92 (<0.01)	<0.01
Absolute native cover (%)	45.83 (18.96)	0.71 (0.14)	33.30 (19.66)	0.49 (0.36)	12.53 (6.92)	0.94 (0.02)	10.76 (8.71)	0.90 (<0.01)	0.01
Absolute noxious cover (%)	2.57 (2.60)	0.90 (0.01)	6.38 (5.09)	0.93 (<0.01)	5.71 (4.32)	0.89 (0.03)	0.94 (1.92)	0.94 (<0.01)	0.01
Relative wetland cover (%)	8.43 (8.77)	0.94 (0.02)	5.93 (3.50)	1.00 (<0.01)	10.15 (8.48)	0.94 (0.02)	5.94 (5.88)	0.95 (<0.01)	0.11
Relative native cover (%)	12.69 (11.83)	0.83 (0.06)	8.91 (4.43)	0.94 0.02)	11.23 (2.19)	0.94 (0.02)	5.02 (5.91)	0.93 (<0.01)	0.18
Mean C	0.29 (0.17)	0.89 (0.03)	0.28 (0.17)	0.83 (0.06)	0.38 (0.23)	0.89 (0.03)	0.12 (0.12)	0.95 (<0.01)	0.45
Mean weighted C	0.62 (0.37)	0.66 (0.18)	0.30 (0.28)	0.94 (0.02)	0.51 (0.24)	0.83 (0.06)	0.22 (0.19)	0.95 (<0.01)	0.12
Shrub richness	[1]Not performed						0.14 (0.36)	0.94 (<0.01)	0.58
Species richness	6.67 (2.58)	0.96 (<0.02)	13.67 (4.93)	0.72 (0.10)	7.67 (4.50)	0.64 (0.17)	4.29 (4.30)	0.88 (<0.01)	<0.01
Herbaceous richness	6.50 (2.66)	1.00 (<0.01)	13.17 (4.26)	0.60 (0.24)	7.33 (3.98)	0.60 (0.24)	3.71 (3.82)	0.86 (<0.01)	<0.01

[1]Not performed due to low number of sites with shrubs

variability was not tested in the LPI method due to the greater time requirement, literature suggests that intercept methods are the most objective (Elzinga and others, 1998; Coulloudon and others, 1999; Vittoz and Guisan, 2007).

Variability Between Visits

Pseudoturnover rates between intra-annual surveys were similar to inter-observer rates, whereas inter-annual rates were more than double intra-annual and inter-observer rates. Such a high rate is likely explained by differences in time and surveyors as all intra-annual and inter-observer sites were surveyed by the same two observers, while inter-annual sites were surveyed by different observer teams and may have been subject to true turnover. Additionally, a one-hour time-limited meander was used in previous years, whereas a progressive timed meander (allowing for survey times greater than 1 hour) was used in 2017. Cases of large intra-annual differences in plant cover were likely caused by species misidentification, while large inter-annual differences were likely caused by a combination of actual change, species misidentification, and species identified to genus only. For example, phenology may play a role in affecting cover estimates for some species within a year, while true expansion of the same species could lead to large changes in cover across years. Discrepancies in cover estimates within and across years were often due to grass and graminoid species, highlighting the challenges associated with estimation and identification of this group of plants. For the inter-annual comparison, there were also large discrepancies in cover estimates among aquatic species, potentially due to changing hydrologic conditions at sites.

We did not find any evidence that surveying sites at different times during the growing season was inappropriate as there were no patterns in vegetation metrics to suggest that surveys were too early or too late in the season. Inter-annual comparisons, despite the larger number of factors that affect vegetation data across years, showed only slightly higher mean differences in metrics than intra-annual comparisons for mean C and relative cover metrics, though absolute cover and richness-based metrics showed much larger differences.

Variability Between Methods

Differences in richness and cover estimates between methods supported results found in similar studies comparing visual and point-intercept survey methods (Kercher and others, 2003; Korb and others, 2003; Godinez-Alvarez and others, 2009). Richness values were higher in the NWCA and meander methods, with the meander detecting the greatest number of species, likely because the method surveyed the largest area. The greater number of species detected by the meander survey may be useful in the detection of rare, sensitive, and noxious species. LPI produced higher absolute cover estimates than both of the other methods. Differences were small for mean C and the relative cover-based metrics and these metrics were highly correlated between methods, suggesting all methods would produce comparable site rankings for these metrics.

Time requirements and ease of use are other important considerations when selecting a survey method. The progressive timed meander stands out in terms of ease of use and suitability for a rapid assessment method for its lower time and personnel requirements involved in setup and flexibility in relation to the complexity of the vegetation community. Furthermore, vegetation metrics from the timed meander were very strongly correlated with metrics calculated from the more intensive NWCA plot data. The timed meander took an average of 50 minutes per site, with the last species recorded between 8 and 188 minutes, depending on the site. The LPI was the most time-consuming to set up and execute and was best done with two people, which further increases the time and resources needed to conduct the survey.

RECOMMENDATIONS AND CONCLUSIONS

While some comparisons in our study are limited by small sample sizes (notably, inter-observer for NWCA plots, intra-annual, and three methods), the results support several recommendations to improve URAP vegetation survey methods. First, mean C and metrics calculated from relative rather than absolute cover in data analysis should be used as these metrics showed greater consistency in all tested methods and time scales. If absolute metrics are to be analyzed, other, more repeatable methods would need to be used for monitoring. For mean C, it is important for surveyors to continue surveying for the full time allotment at species-poor sites because missing species will have a larger impact on mean C at these sites than at species rich-sites. Second, consistency in species identification can be improved by training and retaining qualified botanists and encouraging plant collections in the field. Third, the progressive timed meander survey for vegetation surveys will be used moving forward. The timed meander was the fastest to implement, was strongly correlated to the more intensive NWCA method, captured the most area, and detected the greatest number of species which may be important for detection of noxious weeds, sensitive plant species, and other species of interest.

ACKNOWLEDGMENTS

We appreciate the work of UGS employee Coryna Hebert who assisted with fieldwork for this project. This manuscript was improved by a thoughful review from Rebekah Downard at the Utah Division of Water Quality. This work would not have been possible without the support of the land owners and land managers who allowed us to access their property. Funding for this work was provided by U.S. Environmental Protection Agency Wetland Program Development Grant #CD-96851201.

REFERENCES

Coulloudon, B., Eshelman, K., Gianola, J., Habich, N., Hughes, L., Johnson, C., Pellant, M., Podborny, P., Rasmussen, A., and Robles, B., 1999, Sampling vegetation attributes: U.S. Bureau of Land Management Technical Reference, 1734–4, 163 p.

Elzinga, C.L., Salzer, D.W., and Willoughby, J.W., 1998, Measuring & monitoring plant populations: U.S. Bureau of Land Management, Technical Reference 1730-1, BLM/RS/ST-98/005+1730, 477 p.

Everson, C.S., and Clarke, G.P.Y., 1987, A comparison of six methods of botanical analysis in the montane grasslands of Natal: Vegetatio, v. 73, no. 1, p. 47–51.

Godínez-Alvarez, H., Herrick, J.E., Mattocks, M., Toledo, D., and Van Zee, J., 2009, Comparison of three vegetation monitoring methods: Their relative utility for ecological assessment and monitoring: Ecological Indicators, v. 9, no. 5, p. 1001–1008.

Jones, W.M., 2005, A vegetation index of biotic integrity for small-order streams in southwest Montana and a floristic quality assessment for western Montana wetlands: Helena, Montana Natural Heritage Program, 69 p.

Kercher, S.M., Frieswyk, C.B., and Zedler, J.B., 2003, Effects of sampling teams and estimation methods on the assessment of plant cover: Journal of Vegetation Science, v. 14, no. 6, p. 899–906.

Korb, J.E., Covington, W.W., and Fule, P.Z., 2003, Sampling techniques influence understory plant trajectories after restoration—An example from ponderosa pine restoration: Restoration Ecology, v. 11, no. 4, p. 504–515.

Kuznetsova, A., Brockhoff, P.B., Christensen, R.H.B., 2017, lmerTest Package—Tests in linear mixed effects models: Journal of Statistical Software, v. 82, no. 13, p. 1–26, https://doi.org/10.18637/jss.v082.i13.

Lenth, R., 2019, emmeans—Estimated marginal means, aka least-squares means, R package version 1.3.2. https://CRAN.R-project.org/package=emmeans.

Lepš, J., and Hadincová, V., 1992, How reliable are our vegetation analyses?: Journal of Vegetation Science, v. 3, no. 1, p. 119–124.

Menuz, D., 2016a, Analysis of landscape and habitat variables associated with boreal toad (*Anaxyrus boreas*) breeding locations: Salt Lake City, Utah Geological Survey, contract deliverable for the Utah Department of Natural Resources Endangered Species Mitigation Fund, 39 p.

Menuz, D., 2016b, Development of a watershed stress model for wetlands and other aquatic systems: Salt Lake City, Utah Geological Survey, contract deliverable for the Utah Department of Natural Resources Endangered Species Mitigation Fund, 12 p.

Menuz, D., Sempler, R., and Jones, J., 2016, Weber River Watershed wetland condition assessment: Salt Lake City, Utah Geological Survey, contract deliverable for the U.S. Environmental Protection Agency, 106 p.

Menuz, D., and Sempler, R., 2018, Jordan River Watershed wetland assessment and landscape analysis: Salt Lake City, Utah Geological Survey, contract deliverable for the U.S. Environmental Protection Agency, 82 p.

Minnesota Pollution Control Agency (MPCA), 2014, Rapid floristic quality assessment manual wq-bwm2-02b: Minnesota Pollution Control Agency, St. Paul, MN, p. 1–44.

Morrison, L.W., 2016, Observer error in vegetation surveys—a review: Journal of Plant Ecology, v. 9, no. 4, p. 367-379.

Nilsson, I.N., and Nilsson, S.G., 1985, Experimental estimates of census efficiency and pseudoturnover on islands—error trend and between-observer variation when recording vascular plants: The Journal of Ecology, v.73, no.1, p. 65–70.

Omernik, J.M., 1987, Ecoregions of the conterminous United States: Annals of the Association of American Geographers, v. 77, no. 1, p. 118–125, scale 1:7,500,000.

R Core Team, 2018, R—A language and environment for statistical computing, R Foundation for Statistical Computing: Vienna, Austria, URL https://www.R-project.org/.

Rocchio, J., 2007, Floristic quality assessment indices for Colorado plant communities, Fort Collins, unpublished report prepared for Colorado Department of Natural Resources and U.S. EPA Region 8 by the Colorado Natural Heritage Program, 234 p.

Rocchio, F.J., and Crawford, R.C., 2013, Floristic quality assessment for Washington vegetation: Olympia, Washington Natural Heritage Program, 49 p.

Sykes, J.M., Horrill, A.D., and Mountford, M.D., 1983, Use of visual cover assessments as quantitative estimators of some British woodland taxa: The Journal of Ecology, v. 71, no. 2, p. 437–450.

USEPA, 2016, National wetland condition assessment 2016—field operations manual, EPA-843- R-15-007: U.S. Environmental Protection Agency, Washington, DC.

Vittoz, P., and Guisan, A., 2007, How reliable is the monitoring of permanent vegetation plots? A test with multiple observers: Journal of Vegetation Science, v. 18, no. 3, p. 16.

Washkoviak, L., Heidel, B., and Jones, G., 2015, Floristic quality assessment for Wyoming flora—developing coefficients of conservatism: Lander, Wyoming, Wyoming Natural Diversity Database, prepared for U.S. Army Corps of Engineers agreement number G15AC00484, 19 p.

APPENDICES

Appendix A

Vegetation Survey Field Forms

Timed Meander Datasheet

Site ID:		Observer:	Date:

Identify the number of communities present within the AA. Allow 30 minutes for the first community and add 20 minutes for each additional community. Note exact time that each species was encountered, starting from 0 minutes into the survey. If < 3 new species are encountered during the last 10 minutes, stop meander. If ≥ 3 new species are found during last ten minutes, continue for additional 10 minutes. Continue until < 3 new species are found in time period. Determine % cover of each species in the AA at the end of the meander.

Community 1:____
Community 2:____
Community 3:____
Community 4:____

Initial length of survey: _________ minutes

Height Class **A**: <0.5m **B**: 0.5-1m **C**:1-2m **D** 2-5m **E** 5-10m **F** >10m

Scientific Name/Pseudonym	Code	Time	% Cover	Agreed Cover	Height Class	Collection #	Community?				% Cover in Plot			
							1	2	3	4	1	2	3	4

NWCA Vegetation Plot Survey Form

Site ID:	Plot#	Observer:	Date:
Identify all species and their percent cover of the 10 x 10 plot. Height Class **A**: <0.5m **B**: 0.5-1m **C**:1-2m **D** 2-5m **E** 5-10m **F** >10m			

Scientific Name/Pseudonym	Code	% Cover	Height Class	Collection #

Line-point Intercept Data Form

Site ID:	**Observer:**	**Recorder:**	**Date:**
Transect #:	**Direction:**	**Point Spacing Interval:**	

Directions:
For water, record whether location is dry (D) or saturated (S), or record the water depth in cm.
For soil, record the bottom layer as S = Soil/sand/sediment**,** G = Gravel/cobble (~2–250 mm)**,** R = Rock/boulder/bedrock (>250 mm)**,** A = algae, AC = chara, M = moss
For vegetation, either species code or SDT, SDS, SDH (standing dead tree (>5cm dbh, shrub, herb), include L in canopy layers (surface litter)

Pt.	Top canopy	Lower canopy layers			Water (D, S, #)	Soil (S, G, R, L, A, M)	Pt.	Top canopy	Lower canopy layers			Water (D, S, #)	Soil (S, G, R, L, M)
		1	2	3					1	2	3		
0.5							13						
1							13.5						
1.5							14						
2							14.5						
2.5							15						
3							15.5						
3.5							16						
4							16.5						
4.5							17						
5							17.5						
5.5							18						
6							18.5						
6.5							19						
7							19.5						
7.5							20						
8							20.5						
8.5							21						
9							21.5						
9.5							22						
10							22.5						
10.5							23						
11							23.5						
11.5							24						
12							24.5						
12.5							25						

Top canopy codes: Sp., or none Lower canopy layer codes: Sp., L (herbaceous litter), FWD (Fine woody debris ≤ 2.5cm), CWD (Coarse woody debris > 2.5cm) Standing dead = still attached at base/ Litter = not attached	**Soil Surface:** S = Soil/sand/sediment G = Gravel/cobble (~2–250 mm) R = Rock/boulder/bedrock (>250 mm) DL = litter dense enough to obscure boot RL = remaining litter M = moss

Site Sketch Form

Site ID:	Surveyors:	Date:
Site Sketch: Define scale for grid, add north arrow. Mark inlets and outlet if present in or adjacent to AA.		

Appendix B
Vegetation Condition Metrics

URAP uses qualitative metrics to evaluate wetland condition. Metrics are generally scored by evaluating which of four potential states most closely describes the assessed wetland. States reflect the continuum of potential conditions, from reference standard to highly degraded, that may be found for a particular aspect of wetland condition. States are typically assigned letter ranks from A to D, though the best condition state at some sites is assigned a value of AB because of the difficulty in distinguishing between A and B states. Two vegetation metrics evaluated in this study, absolute noxious cover and relative native cover, are used as condition metrics by URAP (tables B1 and B2). The condition metrics can either be rated using best professional judgement or with calculated values from more intensive plant community data, such as that collected by meander surveys or NWCA plots.

We evaluated the effect of differences between observers, methods, and times of the year on the two vegetation-based condition metrics. We used calculated values for the absolute noxious cover and relative native cover metrics to assign condition metric ranks between A and D and converted the ranks to point values as shown in tables B1 and B2. We assessed variability of metric responses using Krippendorff's α (Kα), a reliability coefficient used to measure agreement among observers, with Kα values of 1 indicating perfect agreement and 0 indicating agreement no better than chance (Krippendorff, 2011). Kα works with two or more raters and with nominal, ordinal, and interval data and can handle incomplete datasets, though Kα penalizes datasets where coders do not vary much in their responses (Krippendorff, 2004). Kα was calculated using the "irr" package (Gamer and others, 2012) in the statistical software R (R Core Team, 2016) with data type "ordinal." We classified Kα scores as indicating poor, slight, moderate, substantial, and near-perfect agreement based on thresholds used in a recent study evaluating the repeatability of the URAP condition method between teams of observers (table B3; Menuz and McCoy-Sulentic, in review).

Table B1. *Metric states for the relative cover of native plant species metric.*

Rank	Value	State
AB	5	Assessment plot contains >95% relative cover of native plant species.
C	3	Assessment plot contains 80–95% relative cover of native plant species.
C-	2	Assessment plot contains 50–80% relative cover of native plant species.
D	1	Assessment plot contains <50% relative cover of native plant species

Table B2. *Metric states for the absolute cover noxious weeds metric.*

Rank	Value	State
A	5	Noxious weeds absent.
B	4	Noxious weeds present, but sporadic (<3% absolute cover).
C	3	Noxious weeds common (3–10% cover).
D	1	Noxious weed abundant (>10%) cover.

Metrics were frequently assigned different values at the same site by different observers, methods, or survey time (table B4). The highest rate of perfect agreement in ratings was between NWCA plots and meander survey, where 20 of 21 sites were scored the same by both observers, and the lowest rate was in the comparison of the three methods, where only one of six sites was scored the same by all three methods. At least two of the three methods and two of the three visits always agreed with one another in the intra-annual and three-method comparison, though the visits and methods that agreed were not consistent across sites.

Metrics were always scored within a rank of one another (e.g., A and B, not A and C) at the same site except in the inter-annual comparison. Large differences in rank in the inter-annual comparison were driven by true site changes over time. For example, three of the four largest differences in the absolute value of noxious weed cover (≥28% cover increases) were due to increases in

Phragmites australis cover, which likely represent true expansion of the species or reestablishment after treatment, based on evidence from site photographs. At another site, the relative native cover rank differed by two ranks at a managed impoundment that went from completely inundated to totally dry from 2015 to 2017. Native cover was dominated by duckweed (a floating aquatic species) during the first visit but the species was not observed during the second survey as there was no standing water at the site.

We found substantial or greater agreement based on Kα values for most comparisons except the inter-observer meander six-site comparison for both metrics and the inter-annual comparison for the absolute noxious cover method (table B4). However, the inter-observer agreement in meander surveys was near perfect when all 19 sites were evaluated, suggesting that the results for the former may be influenced by small sample size. The inter-annual differences likely reflect true changes in wetland condition at sites, as detailed above.

Table B3. *Thresholds used to interpret Krippendorff's α values for observer (or method or time of visit) agreement.*

Krippendorff's α	Agreement Level
>0.80–1	Near perfect
>0.60–0.80	Substantial
>0.40–0.60	Moderate
>0.20–0.40	Slight
≤0.20	Poor

Table B4. *Agreement between surveyors, visits, and methods in condition metric ranks, as measured by Krippendorff α values and percent of site with perfect agreement.*

Comparison	# Sites	Krippendorff α		Percent of Sites with Perfect Agreement	
		Relative native cover	Absolute noxious cover	Relative native cover	Absolute noxious cover
Meander (2 observers)	19	0.86	0.82	78.9	52.6
NWCA plots (2 observers)	6	0.78	0.78	83.3	83.3
Meander (2 observers)	6	0.20	-0.12	50.0	16.7
Intra-annual (3 visits)	6	0.67	0.93	33.3[1]	66.7[1]
Inter-annual (2 visits)	18,19	0.71	0.46	61.1	52.6
NWCA v Meander v LPI (3 methods)	6	0.61	0.77	16.7[1]	16.7[1]
NWCA v Meander (2 methods)	21	0.99	0.95	95.2	95.2

[1]Indicates percent of sites where all three visits or methods agreed.

REFERENCES

Gamer, M., Lemon, J., Fellows, I., and Singh, P., 2012, Irr—various coefficients of Interrater Reliability and Agreement: R package version 0.84, Online, https://CRAN.R-project.org/package=irr.

Krippendorff, K., 2004, Reliability in content analysis−Some common misconceptions and recommendations: Human Communication Research, v. 30, no. 3, p. 411–433.

Krippendorff, K., 2011, Computing Krippendorff's alpha-reliability: Philadelphia, University of Pennsylvania Departmental Papers: Online, https://repository.upenn.edu/asc_papers/43, 10 p.

Menuz, D. and McCoy-Sulentic, M., 2019, Validation of a rapid wetland assessment protocol for Utah−Testing inter-observer and intra-season variability: Salt Lake City, Utah Geological Survey Report of Investigation 278, 9 p., 4 appendices.

R Core Team (2018), R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria, URL https://www.R-project.org/.