Wetlands Ecol Manage (2016) 24:609–622
DOI 10.1007/s11273-016-9491-1

ORIGINAL PAPER

Evaluating a sampling protocol for assessing plant diversity in prairie fens

Rachel A. Hackett · Michael J. Monfils · Anna K. Monfils

Received: 11 January 2016 / Accepted: 4 April 2016 / Published online: 12 April 2016

Abstract Prairie fens are globally vulnerable wetlands that are considered a conservation priority due to threats to their high biodiversity and hydrological functions. Establishing a thorough and repeatable plant sampling protocol is critical to evaluating conservation and management initiatives. Our goal was to evaluate a sample methodology designed to assess prairie fen plant diversity and determine if it produced results (1) representative of site diversity, (2) comparable among fens, and (3) efficient to collect. Nineteen fens between 8.5 and 28.4 ha were surveyed twice within one growing season during 2012 and 2013 field seasons using an area-proportional, random design. The turnover in species between spring and summer sampling periods within a site ranged from 8 to 50 %. Sample coverage of total estimated plant species richness ranged from 84.8 to 95.0 % with a mean of 90.1 %. We compared results from our area-proportional, random design to simulated random samples of 10, 15, 20, 25, 30, 35 and 40 quadrats per site. No significant difference was found in sample coverage per fen when using sampling rates of 25, 30, or 35 quadrats per site versus the area-proportional design. Shannon's diversity index and floristic quality index differed by sample period and number of quadrats sampled per fen. Our sample design produced acceptable levels of coverage and facilitated comparisons across fens. Our methodology could be applied to future research, restoration monitoring, and conservation planning efforts in Midwestern prairie fens.

Keywords Floristic quality · Phenology · Plant diversity · Prairie fen · Sample size · Wetland

Electronic supplementary material The online version of this article (doi:10.1007/s11273-016-9491-1) contains supplementary material, which is available to authorized users.

R. A. Hackett (✉) · A. K. Monfils
Department of Biology, Institute for Great Lakes Research, Central Michigan University, Brooks Hall 217, Mount Pleasant, MI 48859, USA
e-mail: hacke1ra@cmich.edu

M. J. Monfils
Michigan Natural Features Inventory, Michigan State University Extension, P.O. Box 13036, Lansing, MI 48901, USA

Introduction

Prairie fens are sedge-dominated, groundwater-fed wetlands found throughout the glaciated upper Midwest of the United States and southern Ontario, Canada. High calcium and magnesium levels, low nitrogen and phosphorus levels, and the influence of surrounding systems make these ecosystems one of the most species diverse in the temperate region (Moran 1981; Amon et al. 2002; Bedford and Godwin 2003; Rydin and Jeglum 2006). The presence of prairie graminoids and forbs in addition to calcareous and fen plant specialists distinguish prairie fens from other

fens. The high biodiversity of prairie fens, their important wetland functions, and their locally and globally recognized vulnerability marks them as a priority for conservation (Spieles et al. 1999; Bedford and Godwin 2003).

To assess threats and monitor management success in prairie fens, a reliable and consistent method for assessing diversity is needed. Ecological processes rely heavily on plant establishment, vegetative regimes are used to define ecosystems, and vegetation is relatively simple to assess; as such, plant diversity is a commonly used proxy for overall biodiversity and health of an ecosystem (Tilman et al. 1997; Young 2000; Ruiz-Jaén and Aide 2005; Tilman et al. 2014). In Michigan, we have limited records of historical prairie fen plant biodiversity (e.g., Bassett 2004; Crancer 2011; Fiedler and Landis 2012), and no baseline diversity metrics are in place to assess change in these wetland systems. A comprehensive assessment of plant diversity in prairie fens is needed to provide a baseline of plant diversity metrics that can inform future evaluations of ecosystem threats (Spieles et al. 1999). Furthermore, results from multiple projects using a consistent methodology among prairie fens, could provide useful information to inform conservation decisions and evaluate ongoing management efforts.

Prairie fens have several vegetative attributes that are important to consider when establishing a sampling method to determine plant diversity in the field. These wetlands contain both wetland and prairie plant species and are dominated by plant families with different blooming periods (Moran 1981; Carpenter 1995; Spieles et al. 1999). Such phenology can affect sample representation (Lopez et al. 2002; Matthews 2003; Tucker et al. 2005). Many cryptic, rare, and sensitive species also occur in prairie fens (Moran 1981; Spieles et al. 1999; Amon et al. 2002; Bedford and Godwin 2003; Kost and Hyde 2009). Prairie fens are structurally variable, having up to four vegetation zones (i.e., sedge meadow, inundated flat, calcareous groundwater seep, wooded fen) that vary in size and distribution (Spieles et al. 1999). These characteristics may cause factors such as sampling period and sampling size to limit the ability to capture representative samples, high percentage of the species diversity, and comparable samples effectively and efficiently.

For plant diversity metrics to be used efficiently and consistently across prairie fens, a sample methodology must be established that (1) samples sufficient site diversity, (2) produces results that are comparable among fens, and (3) is efficient to implement. Our goal was to develop a sampling protocol that meets these requirements that could be used consistently among fens by researchers and conservationists. We examined data from a field study of plant biodiversity in 19 Michigan prairie fens to evaluate our sampling methodology to ensure that it accounts for the effects of sampling period and sample size on plant composition and commonly used diversity measures. As part of our evaluation, the results of our area-proportional methodology were compared with simulated results as if we had used a "standard" sampling procedure with equal quadrat replication across all fens (e.g., 20 quadrats at every site). Results from our analyses could inform the development of an efficient sample design and protocol that could be applied to future studies, restoration projects, and conservation plans being implemented in Midwestern prairie fens.

Methods

Site selection

The State of Michigan, USA, is in the north-central portion of the prairie fen range. The Michigan Natural Features Inventory (MNFI) monitors over 150 occurrences of prairie fens in Michigan (Michigan Natural Features Inventory 2011). Glacial geology and hydrology limits the range of prairie fens to the southern half of the Lower Peninsula of Michigan, with one group almost entirely contained in the two ecoregions called the Ann Arbor Moraine Ecoregion and Jackson Interlobate Ecoregion and a second group in the western Lower Peninsula crossing seven ecoregions (Fig. 1; Albert 1995). This study focused on prairie fens in the Ann Arbor Moraine Ecoregion and Jackson Interlobate Ecoregion, because these sites represented well characterized fens of high conservation concern that spanned a range of conditions, from public to private ownership, intensively managed and unmanaged, and high to low quality. The fens studied were located between 41°45′N and 42°52′N latitude and 83°11′W and 84°58′W longitude.

The Ann Arbor Moraine Ecoregion and Jackson Interlobate Ecoregion range in elevation from 228 to 390 m above sea level, annual precipitation ranges

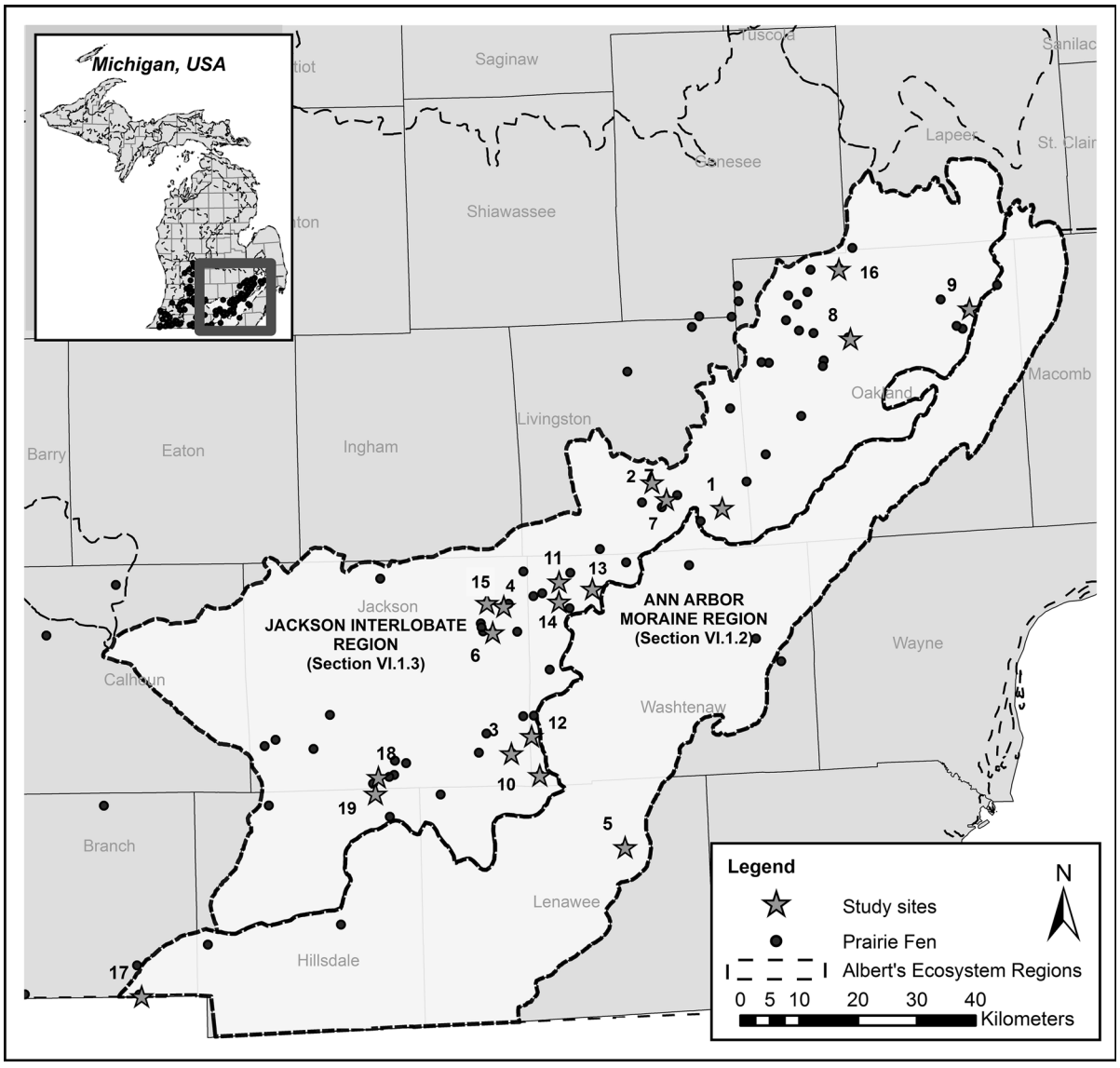

Fig. 1 A map of the Jackson Interlobate and Ann Arbor Moraine Ecoregions in Michigan, USA, marking the nineteen prairie fen study sites with stars (Albert 1995). The *inset map* shows the location of the Jackson Interlobate and Ann Arbor Ecoregions in Michigan, USA. See Table 1 for the prairie fen corresponding to the number listed

from 76 to 91 cm, and extreme minimum air temperature ranges from −33.6 to −30.3 °C. The growing season in the two regions ranges from 140 to 160 days (Albert 1995).

To reduce species-area relationship effects, sites were selected from the same mathematically derived size class using Jenks natural breaks algorithm (Jenks 1967). From the resulting classes, the size class from 8.5 to 28.4 ha was selected to reduce logistical constraints of large sites, reduce diversity limitations of small sites, and include sites of high conservation concern. Ten sites, randomly selected from the 19, were sampled in 2012; the nine remaining sites were sampled in 2013 (Table 1).

Prairie fen delineation

The study site perimeter for each study site was determined using a GIS shapefile provided by MNFI (Michigan Natural Features Inventory 2011), National

Table 1 List of prairie fens sampled in Michigan's Jackson Interlobate and Ann Arbor Moraine Ecoregions as described by Albert (1995)

Order	Site name	Site code	County	Area (ha)
1	Island Lake Recreation Area—Spring Mill Creek Fen	SMC	Livingston	15.6
2	Chilson Fen	CFQ	Livingston	19.4
3	Fay Lake Fen	FLQ	Jackson	19.2
4	Nirdlinger Fen	NFQ	Jackson	11.5
5	Ives Road Fen	IRQ	Lenawee	22.3
6	Waterloo Recreation Area—Mt. Hope Road Fen	MTH	Jackson	10.6
7	Brighton Recreation Area—Little Appleton Lake	LAL	Livingston	11.0
8	Bridge Valley	BVF	Oakland	11.1
9	Bald Mountain Recreation Area—Chamberlain Lakes	CHL	Oakland	12.2
10	Whelan Lake Fen	WLQ	Washtenaw	17.1
11	Park Lyndon Fen	PLF	Washtenaw	25.9
12	Buss Road	BUS	Washtenaw	10.7
13	Riker Lake Prairie Fen	RLF	Washtenaw	17.5
14	Waterloo Long Lake Fen	WLL	Washtenaw	17.0
15	Little Portage Lake Fen	LPL	Jackson	28.4
16	Holly Recreation Area—Brandt Road Fen	BRD	Oakland	8.5
17	Little Fawn River	LFR	Branch	11.4
18	Liberty Fen North	LFN	Jackson	11.6
19	Liberty Fen South	LFS	Hillsdale	16.6

Site order was determined using a random number generator. Site name reflects the name granted to the site from MNFI Element Occurrence database (Michigan Natural Features Inventory 2011). Site code is an abbreviation based on the site name. All counties are in Michigan, USA. The area (ha) reported is based on adjusted perimeters derived from the MNFI database, NAIP digital orthogonal photographs, and ground-truthing

Agriculture Imagery Program (NAIP) 2009 and 2012 digital orthogonal photographs, and ground-truthing. Site perimeters from the MNFI element occurrence shapefile were updated with the NAIP digital orthogonal photographs to match current vegetation, most often removing areas of recent canopy closure. Areas were ground-truthed during the spring sampling period where the distance was greater than 100 m between the perimeter in the MNFI shapefile and the perimeter updated with the NAIP photographs. The ground-truthed areas were examined against the criteria of a prairie fen as defined by Spieles et al. (1999) and Kost et al. (2007): areas contained wetland soils, showed signs of at least seasonal saturation, were composed of less than 25 % tree canopy cover or 50 % shrub canopy cover, and contained at least two indicator species [e.g., *Larix laricina* (Du Roi) K. Koch, *Toxicodendron vernix* (L.) Kuntze, *Parnassia glauca* Raf., *Dasiphora fruticosa* (L.) Rydb., *Pycnanthemum virginianum* (L.) B.L.Rob. & Fernald, *Solidago ohioensis* Riddell, *S. riddellii* Frank, *Sorghastrum nutans* (L.) Nash]. Areas not fitting the prairie fen criteria were removed from the site perimeter. Updated perimeters were used for sampling quadrat placement and analyses.

Sampling methodology

We used an area-proportional, random sampling methodology similar to the those of Sluis (2002) and Johnston et al. (2009) to give equal opportunity to capture each of the four patchy vegetation zones; our method differed in the length of the segments used to divide the baseline and transects and the minimum number of quadrats sampled per site (Fig. 2). At each site, a baseline was drawn across the longest portion of the adjusted perimeter in ArcMap (version 10.1, Environmental Systems Research Institute, Redlands,

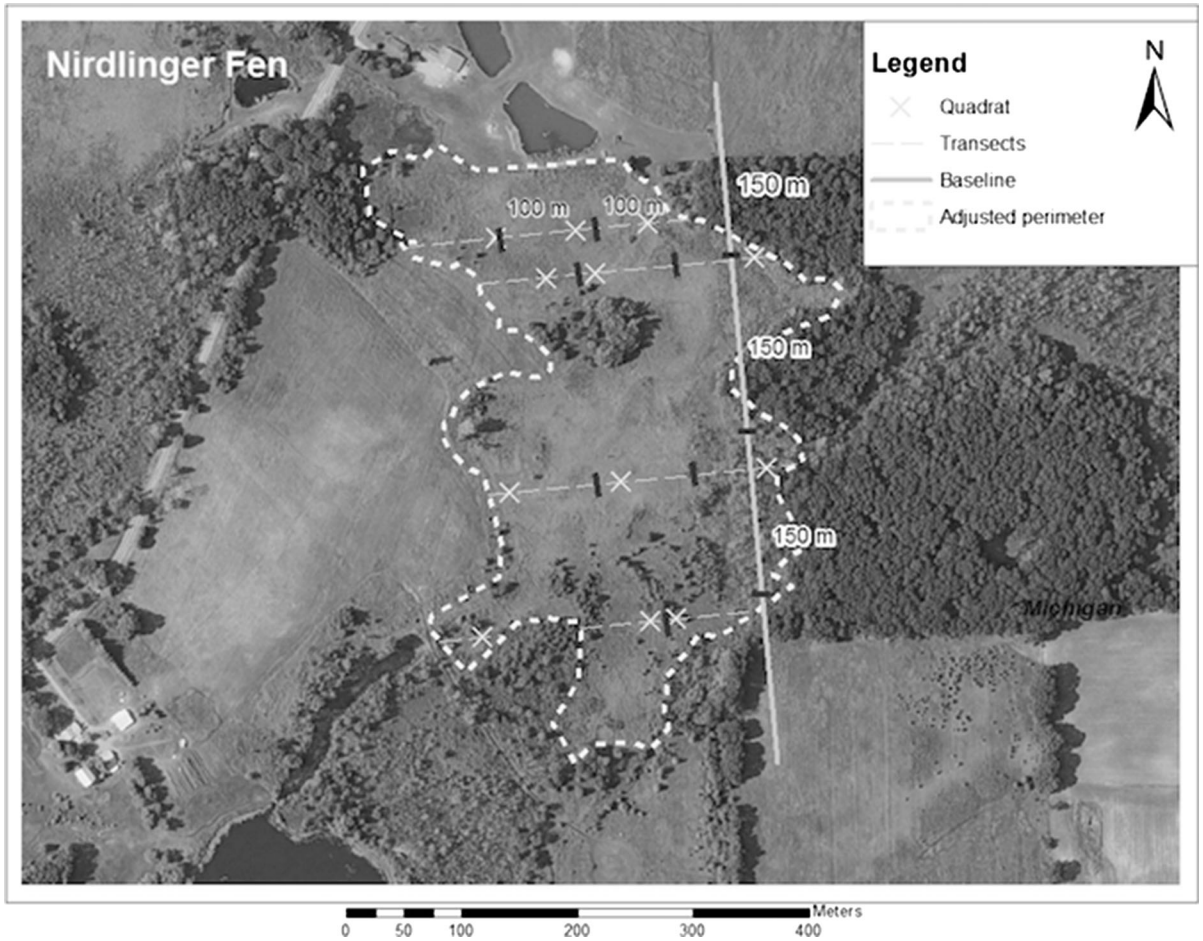

Fig. 2 Random, area-proportional sampling method used at each site. The *thick, dashed line* is the perimeter, the *solid line* is the *baseline*, the *thin, dashed lines* are transects, and "x" represent quadrats sampled

CA). For each sampling period, one transect was placed at a random point for every 150 m of the baseline and extended perpendicular to the baseline for the width of the site (e.g., a 750 m baseline had five transects). For every 100 m of each transect, one 1 m^2 quadrat was placed at a random point (e.g., a 300 m transect had three quadrats). A minimum of ten quadrats was sampled at each site per sampling period regardless of prairie fen dimensions, for a minimum of 20 per site per growing season. Gotelli and Colwell (2011) suggested a minimum of 20 quadrats per site is required to employ post hoc interpolation methods for adjusting species richness.

On four occasions at three sites (i.e., BRD, BUS, BVF), less than ten quadrats were generated with the procedure above. For these occasions, a tenth quadrat was determined in the field. The tenth quadrat was placed at a random bearing and, on that bearing, a random distance less than 200 m from quadrat one within the perimeter of the site.

If a quadrat fell outside of prairie fen community or was determined unsuitable for sampling (e.g., shrub-carr, upland, open water), one of two adjustments were made to the sample design: (1) if prairie fen was less than 100 m from the original quadrat coordinates, a replacement quadrat was set at a random point at less than 25 m along a bearing set perpendicular to closest border to the original coordinates, or (2) if the prairie fen was greater than 100 m from the original quadrat coordinates, the quadrat was removed without replacement. Where applicable, the site perimeter was adjusted accordingly before the next sampling period.

This method resulted in a mean sampling rate of one quadrat per 1 ha of prairie fen per sampling period.

Vegetative sampling

Vegetation sampling at each study site was performed twice in a growing season: spring (May 14, 2012, to June 29, 2012; May 23, 2013, to July 3, 2013) and summer (July 23, 2012, to August 29, 2012; August 5, 2013, to September 6, 2013). In each 1 m^2 quadrat, vascular plants were identified to lowest taxonomic unit and visually assigned a cover class using Daubenmire's cover class scheme (Daubenmire 1959). A voucher specimen was collected for each species encountered over the course of this study. Additionally, if a species could not be identified in the field, a specimen was collected for later identification. Voucher specimens for *Toxicodendron radicans* (L.) Kuntze (poison ivy) and *T. vernix* (poison sumac) were not collected for safety reasons. All voucher specimens were deposited in the Central Michigan University Herbarium (CMC) and images and associated data are available at midwestherbaria.org (Online Resource 1).

Data analysis

All data analyses were run using the statistical program R (version 3.1.1, R Foundation for Statistical Computing, Vienna, Austria) unless otherwise specified.

Plant diversity measures

For each study site, the species and cover class data were used to calculate total, native, and exotic species richness (S, S_{native}, S_{exotic}; Herman et al. 2001; Reznicek et al. 2011); exotic relative abundance; Shannon's Diversity Index (H'; Shannon and Weaver 1949); normal and adjusted mean coefficient of conservatism ($\bar{C}$; Swink and Wilhelm 1994; Wilhelm and Masters 1995; Herman et al. 2001; Reznicek et al. 2011); weighted coefficient of conservatism (wC; Bourdaghs 2012; MPCA 2014); and normal and adjusted Floristic Quality Index (FQI; Herman et al. 2001). The adjusted $\bar{C}$, wC, and adjusted FQI took exotic species into account by assigning them a C of zero. The measure wC, unlike other C measures, incorporates the relative abundance of individual species, and has been shown to be a more responsive indicator than $\bar{C}$ (Bourdaghs 2012; MPCA 2014).

Taxa identified to species or genus level were included in calculation of S and H', and only those identified to species were included in the calculation of S_{native}, S_{exotic}, normal and adjusted $\bar{C}$, wC, and FQI (Online Resource 2). Plant diversity measures were calculated for the following three datasets: spring sampling period, summer sampling period, and the pooled data for the entire growing season.

The normal and adjusted values of $\bar{C}$ and those of FQI were strongly correlated (Pearson's coefficients = 0.93, 1.00, respectively; p-values <0.001); therefore, adjusted values were not reported in further analyses. Exotic relative abundance, which was expressed as a proportion, was arcsine-square root transformed in further analyses.

Sample period effects

To determine if spring and summer sample periods are both necessary to collect a sample representative of the site, species composition between sampling periods was compared. Percent turnover of species derived from numbers equivalent beta (β)-diversity was chosen over other similarity indices [e.g., Sørensen (Sørensen 1948); Jaccard (Jaccard 1912); Horn-Morisita (Horn 1966)], because it is more easily interpreted and compared than similarity indices (Jost 2006, 2007). The percent turnover was calculated for orders zero, one, and two using the R "vegetarian" package with no weights between sampling periods for each site (Jost 2006, 2007; Charney and Record 2009). Order zero uses species presence and absence and gives equal weight to all species, similar to the Sørensen index (Sørensen 1948). Order one weighs species by their relative abundance, like Shannon's entropy (Shannon and Weaver 1949). Order two applies a greater emphasis on dominant species, similar to the Horn-Morisita index (Horn 1966).

Sample period has been shown to influence diversity measures, especially FQI (Francis et al. 2000; Matthews 2003; Matthews et al. 2005). To compare the effect of sample period on all diversity measures across all sites, a paired Hotelling's T^2 test was conducted with all diversity measures that had normally distributed differences between sample periods. To compare the effect of sample period on

each diversity measure individually, a paired *t* test was used for diversity measures with normally distributed differences; a Wilcoxon paired signed-rank analysis was used for diversity measures with non-normally distributed differences. To explore patterns between diversity measures calculated in different sample periods, a Pearson's correlation coefficients analysis was conducted. When data were non-normally distributed, a Spearman's rank correlation coefficient analysis was conducted.

Sample coverage

An optimal sample design will capture a high percentage and consistent sample completeness (i.e., proportion of the diversity sampled) across all sites examined (Gotelli and Colwell 2001; Colwell et al. 2012; Chao and Jost 2012). Sample coverage, the percentage of total estimated species richness in a sample, was used to measure sample completeness. To calculate sample coverage at each site, species incidence (i.e., number of quadrats in which a species was detected at an individual site) was derived and imported into the iNEXT program to interpolate and extrapolate sample coverage-area curves (Chao and Jost 2012; Hsieh et al. 2013). From the sample coverage-area curves, sample coverage was recorded for each site at the actual number of quadrats sampled. The minimum, mean, range, and standard deviation of sample coverage was calculated for the area-proportional sampling.

To determine the efficiency of the sample design compared to "standard" methods of sampling (i.e., equal number of quadrats sampled per site, regardless of site size), sample coverage was determined at increments of 10, 15, 20, 25, 30, 35, and 40 quadrats for each site using the sample coverage-area curves. The minimum, mean, range, and standard deviation of sample coverage was calculated across all sites for each increment of quadrat sampling. An analysis of variance (ANOVA) and Tukey's Honest Significant Difference (HSD) was conducted to determine if the sample coverage was significantly different among and between sample sizes.

Performance compared to simulated standard design

Sample size has been shown to affect diversity indices such as *H'* and FQI (Wolda 1981; Francis et al. 2000; Bourdaghs et al. 2006). Shannon's Diversity Index, exotic relative abundance, $\bar{C}$, *wC*, and FQI were compared between the area-proportional sampling and those calculated from simulated standard sampling (i.e., set number of random quadrats sampled at each site). For each site, sample sizes were set at 10, 15, 20, 25, 30, 35, and 40 quadrats and resampled from the quadrat-species dataset 50 times per site at random with replacement. Shannon's Diversity Index, exotic relative abundance, $\bar{C}$, *wC*, and FQI were calculated for each resampled dataset. If the number of quadrats resampled exceeded the actual number of quadrats sampled at a site, the diversity measures were not calculated for that increment of sample effort. Root mean square error (RMSE) was used to quantify differences in diversity measures between area-proportional and the simulated standard design at different quadrat increments.

Results

Area-proportional sampling resulted in 561 quadrats sampled across 19 sites; the mean was 29.5 quadrats per site and the maximum was 53 quadrats at a single site. One hundred and twenty-five days were spent sampling over a 2 year period. Species richness per quadrat ranged from 1 to 30 species and had a mean and median of 14 (SD = 5). Within a single site, total species richness averaged 106 species (SD = 24; Table 2), and gamma species richness among all sites was 299 species (Online Resource 1, 2). Approximately 13 % of the quadrats were moved in field because the quadrat locality was unsuitable for sampling or to be consistent with the definition of a prairie fen (Spieles et al. 1999; Kost et al. 2007).

There were 1057 voucher specimens collected: 997 specimens were identified to 297 different species (*Toxicodendron radicans* and *T. vernix* were not collected), 52 specimens were identified to genus, four to family, and four were unidentifiable (Online Resource 1). Of the 297 species identified, 158 were forbs, 70 were grasses or sedges, 50 were shrubs or trees, 11 were vines, and eight were ferns or fern allies. Three species were of special status: *Arnoglossum plantagineum* Raf. (Michigan special concern), *Carex trichocarpa* Muhl. ex Willd. (Michigan special concern), and *Muhlenbergia richardsonis* Rydb. (Michigan threatened; Reznicek et al. 2011). Twenty-four exotic species were identified.

Table 2 Means of diversity measures by sample period

Period	S	S_{native}	S_{exotic}	H'	ERA (%)	$\bar{C}$	wC	FQI
Spring	74 (18)	68 (16)	5 (3)	3.40 (0.40)	5.4 (3.8)	4.7 (0.3)	4.4 (0.6)	38.7 (6.0)
Summer	81 (21)	75 (19)	5 (3)	3.55 (0.34)	7.2 (5.6)	4.9 (0.4)	4.5 (0.9)	42.1 (6.1)
Both	106 (24)	98 (21)	7 (4)	3.69 (0.35)	6.3 (4.5)	4.9 (0.3)	4.5 (0.6)	47.9 (5.9)

The standard deviation is listed in parentheses. Plant diversity measures are total species richness (S), native species richness (S_{native}), exotic species richness (S_{exotic}), exotic relative abundance (ERA), Shannon's Diversity Index (H'), mean coefficient of conservatism ($\bar{C}$), weighted coefficient of conservatism (wC), and Floristic Quality Index (FQI)

Table 3 Numbers equivalent turnover of orders zero, one, and two between the two sampling periods (spring and summer) at each site (Jost 2007; Charney and Record 2009)

Site code	Numbers equivalent turnover (%)		
	Order 0	Order 1	Order 2
BRD	35	25	18
BUS	30	18	12
BVF	50	26	19
CFQ	37	24	20
CHL	45	27	18
FLQ	38	26	19
IRQ	39	32	33
LAL	42	22	17
LFN	31	22	15
LFR	47	29	18
LFS	45	18	11
LPL	29	29	40
MTH	47	25	14
NFQ	39	29	17
PLF	29	16	9
RLF	37	22	21
SMC	45	22	11
WLL	37	17	10
WLQ	36	21	10
Mean	39	24	18
SD	7	5	8
Minimum	29	16	9
Maximum	50	32	40
Range	21	16	31

The site codes are abbreviations for each site that can be found in Table 1

Sample period effects

The turnover in species between the spring and summer sampling periods within a site ranged from 9 to 50 % across all orders (Table 3). The mean turnover for orders zero, one, and two was 39 % (SD = 7 %), 24 % (SD = 5 %), and 18 % (SD = 8 %), respectively. Across all sites, 34 plant species were detected only in the spring sample period and 53 plant species were detected only in the summer sample period.

The multivariate comparison of diversity measures indicated a significant difference between spring and summer sampling periods (paired Hotelling's $T^2 = 2.95$, p-value = 0.05). When diversity measures were compared individually, t-tests indicated that S, S_{native}, H' and FQI differed by sampling period (Table 4). Exotic relative abundance, $\bar{C}$, and wC were similar between sampling periods (Table 4).

Exotic species richness did not have normally distributed differences, so it was not included in the Hotelling's T^2 statistic and the Wilcoxon paired signed-rank test was used to compare between sampling periods. Exotic species richness was not significantly different between sample periods (Table 4).

All plant diversity measures were significantly correlated (p-value <0.05) between sampling periods except for S_{exotic} and $\bar{C}$ (Table 4). Total species richness, S_{native}, exotic relative abundance, H', and FQI were highly correlated (r >0.70). Weighted coefficient of conservatism was moderately correlated between periods (r = 0.50).

Sample coverage

The sample coverage of the area-proportional sampling methodology ranged from 84.8 to 95.0 % with a mean of 90.1 % (SD 2.7 %; Fig. 3). The minimum, mean, range, and standard deviation of the sample coverage at 10, 15, 20, 25, 30, 35, and 40 sample sizes is also displayed in Fig. 3 and Table 5. The sample coverage of differing sample sizes were significantly different overall (ANOVA F-value 65.1; p-value <0.001). There were no significant differences in sample coverage in pairwise

Table 4 Statistics for the paired difference statistical tests and correlation analyses

Statistic	S	S_{native}	S^{a}_{exotic}	ERA[b]	H'	$\bar{C}$	wC	FQI
$t_{0.05,\ 18}$	−2.4	−2.7	73[c]	−1.75	−2.9	−1.94	−0.42	−3.1
Mean difference	−6.8	−6.6	–[c]	−0.03	−0.15	−0.19	−0.07	−3.4
Paired t-test p-value	0.02	0.02	0.8[c]	0.09	0.01	0.07	0.7	<0.01
Correlation coefficient	0.81	0.81	0.65[d]	0.75	0.82	0.41	0.50	0.70
Correlation p-value	<0.01	<0.01	0.09[d]	<0.01	<0.01	0.08	0.03	<0.01

A paired t-test was conducted for plant diversity measures with normally distributed differences between sampling periods. Pearson's correlation coefficient analysis was conducted for normally distributed measures. For the correlation analysis, all p-values were less than 0.05 except for the noted exception. Plant diversity measures are total species richness (S), native species richness (S_{native}), exotic species richness (S_{exotic}), exotic relative abundance (ERA), Shannon's Diversity Index (H'), mean coefficient of conservatism ($\bar{C}$), weighted coefficient of conservatism (wC), and Floristic Quality Index (FQI)

[a] Non-normally distributed

[b] Exotic relative abundance was arcsine-square root transformed prior to analysis

[c] Wilcoxon paired signed-rank test used instead of paired t-test; the V-statistic is listed instead of the t-statistic

[d] Spearman's rank correlation coefficient used instead of Pearson's coefficient

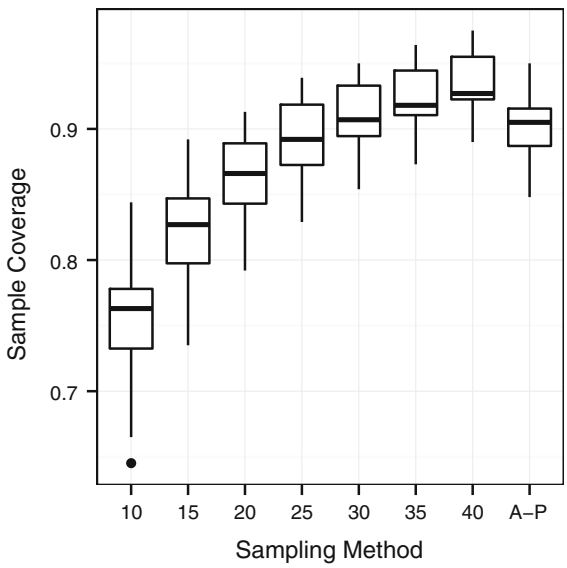

Fig. 3 Distribution of the sample coverage among sites using our area-proportional methodology and simulated "standard" sampling. Simulated samples with equal quadrat replication at all sites are represented in the graph by the number of quadrats sampled per site (e.g., 10, 15, 20) and the area-proportional sampling design is represented by *A-P*

comparisons between the area-proportional design and simulated sampling of 25, 30, or 35 quadrats per site (Tukey's HSD = −0.00947, 0.00879, 0.0237; p-values = 0.987, 0.991, 0.357, respectively).

Performance compared to simulated standard design

For H', the RMSE between the simulated standard sampling at different quadrat increments and the area-proportional sampling decreased with an increase in the number of quadrats sampled per site with sampling at 40 quadrats per site having the least RMSE (0.117; Fig. 4). Simulated standard sampling with less than 35 quadrats per site produced lower H' than area-proportional sampling, especially for sites with greater diversity (Online Resource 3).

For FQI, the RMSE between simulated sampling and the area-proportional design decreased with an increase in the number of quadrats sampled per site, with sampling at 40 quadrats per site having the least RMSE (1.70; Fig. 4). Simulated sampling with less than 35 quadrats resulted in lower FQI compared to the area-proportional design (Online Resource 3).

No trends or patterns were observed for exotic relative abundance, $\bar{C}$, or wC between the area-proportional design and simulated standard sampling (Online Resource 3). There was a decrease in RMSE as simulated standard sampling size increased, but it is likely due to subsampling from a limited number of samples (Fig. 4). Although measured on the same scale (0–10), wC had RMSE twice as large as $\bar{C}$ illustrating a greater variability in the wC measure.

Table 5 Sample coverage by sites for the area-proportional design and simulated sampling using several increments of quadrats per site

Site code	Area in ha (quadrats sampled)	Sample coverage (%)							
		Area-proportional	Quadrat per site increments						
			10	15	20	25	30	35	40
BRD	8.5 (20)	91.1	82.3	88.4	91.1	93.0	94.4	95.6	96.5
BUS	10.7 (21)	91.5	**84.4**	**89.2**	**91.3**	92.5	93.6	94.5	95.3
BVF	11.1 (21)	**84.8**	73.0	80.0	84.2	87.3	89.8	91.8	91.8
CFQ	19.4 (34)	91.6	74.9	82.7	86.8	89.2	90.7	91.7	92.6
CHL	12.2 (31)	85.8	**64.5**	**73.5**	**79.2**	**82.9**	**85.4**	**87.3**	**89.0**
FLQ	19.2 (35)	91.3	70.9	79.1	84.0	87.2	89.5	91.3	92.7
IRQ	22.3 (43)	90.5	66.5	75.1	80.3	83.7	86.2	88.1	89.7
LAL	11.0 (24)	88.9	73.5	81.2	86.1	89.5	92.1	94.0	95.5
LFN	11.6 (27)	92.2	78.5	85.7	89.3	**93.9**	93.1	94.4	95.5
LFR	11.4 (20)	89.2	77.3	84.7	89.2	92.3	94.5	96.1	97.2
LFS	16.6 (25)	88.5	77.2	83.1	86.6	88.5	89.9	91.1	92.2
LPL	28.4 (41)	94.7	78.1	84.7	88.4	90.8	92.4	93.6	94.5
MTH	10.6 (23)	86.6	76.5	82.1	85.2	87.5	89.4	91.0	92.4
NFQ	11.5 (22)	89.8	76.3	84.1	88.6	91.4	93.5	95.1	96.3
PLF	25.9 (53)	**95.0**	77.5	84.1	87.7	90.0	91.6	92.7	93.6
RLF	17.5 (23)	91.9	78.4	85.7	90.0	92.9	**95.0**	**96.4**	**97.5**
SMC	15.6 (34)	89.0	75.1	81.1	84.8	87.2	88.9	90.2	91.4
WLL	17.0 (29)	88.3	73.6	79.5	83.5	86.4	88.8	90.8	92.4
WLQ	17.1 (35)	91.1	70.2	79.2	84.4	87.6	89.7	91.1	92.3
Mean		90.1	75.2	82.3	86.4	89.1	91.0	92.5	93.6
SD		2.7	4.9	4.0	3.4	3.1	2.7	2.6	2.4
Range		10.2	19.9	15.7	12.1	11.0	9.6	9.1	8.5
No. of quadrats		561	190	285	380	475	570	665	760

Values were interpolated and extrapolated using iNEXT program based on species incidences (Chao and Jost 2012; Hsieh et al. 2013). The maximum and minimum of each sampling methodology is bolded. The site codes are abbreviations for each site that can be found in Table 1. Mean, standard deviation (SD), range of sample coverage among sites, and number of quadrats needed to sample 19 sites at various sample sites is listed at the bottom of the table

Discussion

Sample period effects

The growing season is less than half of the calendar year in the Ann Arbor Moraine Ecoregion and Jackson Interlobate Ecoregion and throughout much of the rest of the glaciated upper Midwest. This limited time and resources can restrict researchers' ability to sample multiple fens within multiple sample periods. There is a temptation to sample different prairie fens throughout and across the growing season. When sampling in spring and summer at the same site, there was a pronounced turnover in species, which was particularly strong for order zero. Rare and cryptic species influence order zero more than in other orders. This is consistent with what we know about prairie fens and the high number of rare species that occupy these ecosystems. We have strong support for multi-season sampling in an effort to account for overall species diversity of a site. A sample design including multi-period sampling is best suited for capturing the highest level of diversity and consistently comparing diversity among sites. If multi-period sampling is unfeasible, measures should not be directly compared between sites sampled in different periods with the exception of

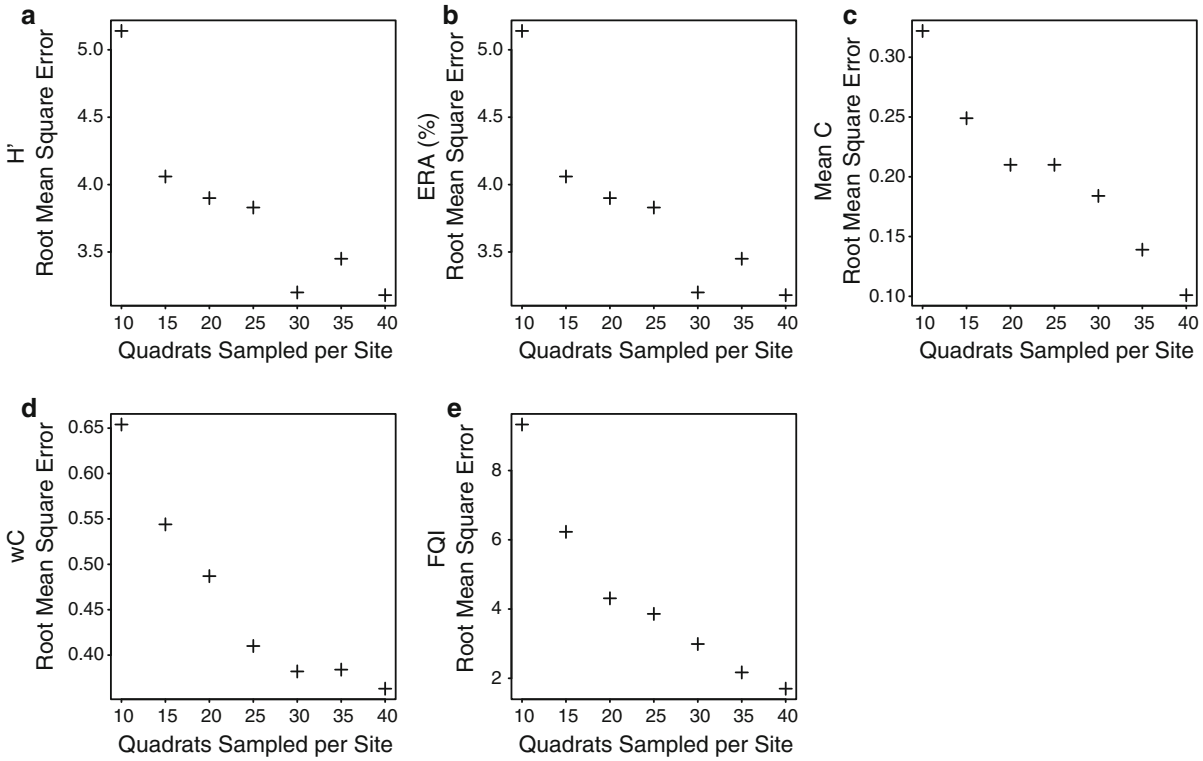

Fig. 4 Scatterplots of the root mean *square error* (RMSE) for simulations of sampling methodologies for diversity measures (**a**–**e**). The *x-axis* is the number of quadrats sampled per site for the simulated standard designs. The "+" *mark* the RMSE between the area-proportional sampling design and the simulated standard sample design

$\bar{C}$ and *wC*. Comparisons of measures such as *S* and *H'* among sites sampled in different periods of a given year may spawn spurious results.

Sample coverage

The vulnerable status of prairie fens and the desire to manage and monitor these ecosystems, calls for a method of assessment that is comparable among sites. Comparing sites at a high and equal sample completeness (e.g., sample coverage) is the best representation of sample comparability across sites (Alroy 2010a, b; Jost 2010; Chao and Jost 2012). A common method used to gather samples of equal sample completeness is a "stop rule" to estimate sample coverage while in the field (Rasmussen and Starr 1979; Chao and Jost 2012). Unfortunately, the area-proportional design was not conducive to using sampling stop rules, because entire sections of the site may not be reached after sampling a small number of sampling units (e.g., quadrats) in the same vegetative zone. As a result larger quadrat numbers must be sampled to minimize differences in sample coverage among sites (Hill et al. 1994). Large sample sizes gathered at all sites may not be the most efficient use of resources and time, when one considers area-species relationships and sampling sites of varying sizes. In prairie fens, the most efficient sampling methodology will capture high sample coverage, is comparable in sample coverage and diversity measures among sites, and will minimize oversampling.

In this study the use of an area-proportional sampling design was explored, and we found this method was able to capture high sample coverage (at least 84.8 %; Fig. 4) with small deviation among sites (SD 2.4 %). The sample coverage-area curves indicated that to reach sample coverage of at least 84.8 % (the minimum coverage determined using the area-proportional sampling design), no individual site needed more than 30 quadrats sampled. If the 84.8 % coverage is acceptable for comparisons, the area-proportional method "over"-sampled at 8 of the

19 sites. The area-proportional method could be more efficient for the fen size range we sampled if an upper-limit of 30 quadrats was applied, reducing the number of quadrats sampled (495; 66 fewer than sampled). This could reduce the time spent in the field by 13 days at the rate of 5 quadrats per day. If an 84.8 % coverage is sufficient and the researcher is concerned with only overall diversity, upper limits for the sample sizes included in this study could be reduced and provide an opportunity to measure additional fens during the growing season.

Performance compared to simulated standard design

Sample methodology and size has been shown to affect diversity measures (e.g., FQI) when not accounting for sample completeness (Wolda 1981; Francis et al. 2000; Bourdaghs et al. 2006). Total species richness is easily adjusted mathematically to account for sample completeness (e.g. Connor and McCoy 1979; Colwell and Coddington 1994; Colwell et al. 2004; Gotelli and Colwell 2011; Chao and Jost 2012). Other diversity measures are not easily adjusted mathematically: $\bar{C}$ and FQI require species specificity, *H'* requires individual species abundances, and exotic relative abundance and *wC* requires both for calculations.

In agreement with other studies (Wolda 1981; Francis et al. 2000; Matthews 2003; Matthews et al. 2005; Bourdaghs et al. 2006; Bourdaghs 2012), reduced sample sizes compromises the use of diversity measures with the possible exceptions of $\bar{C}$ and *wC*. This reemphasizes the need for the development of a standard sampling methodology for prairie fens, and cautions the comparison of most diversity measures among studies unless the sample coverage captured of the sites is comparable (as was the case using our study design). Meta-analyses of prairie fen diversity studies must account for inconsistent sample coverage and sample size in addition to sampling period for results to be valid.

In the simulated sample designs, it must be noted that some of the decrease in variation was likely due to resampling of a limited number of samples, but visible trends, especially at lower quadrat increments, are supported by similar results from other studies (Francis et al. 2000; Matthews 2003).

It should be noted that it was only the effects of sampling size and season on diversity measures that were tested here. These limitations can affect sampling strategies and comparison among sites. The ability of the measures as an indicator of wetland condition or integrity was not tested.

Summary

The sample design employed was consistent and sampled a high percentage of the overall diversity allowing for comparisons among sites. This is a compelling argument for both the sample design and number of quadrats sampled. The method employed was effective at gathering representative diversity that was comparable across fen populations.

Overall, our area-proportional design combined with the intensive sampling (a mean of one quadrat per 1 ha per sample period) resulted in the detection of a large and consistent proportion of the plant species across fens in the size class studied. The high level of turnover detected in species across the season further supports that the phenology of prairie fens plants requires samples to be taken in both the spring and summer to capture representative compositional diversity. The two season sampling protocol must be consistently applied to avoid the possibility of skewed diversity measures and erroneous comparisons that may be produced when samples are taken in different sampling periods.

The sample design employed could be modified to reduce time in the field allow for inclusion of additional sites. Sampling with a maximum of 30 quadrats per site would maintain an average level of sample coverage across sites similar to the sample design employed for this project, with no site being below 84.8 % coverage, saving approximately 13 days of field sampling. We did see a noticeable difference in *H'* and FQI between the area-proportional design and simulated sampling at 25, 30 and 35 quadrats per site. The RMSE consistently becomes smaller as sample number increases, indicating that smaller sample sizes would not reach a consistent and comparable level of detection at lower sample numbers. A reduced number of quadrats would compromise the ability to compare diversity metrics (other than $\bar{C}$) across sites.

The patterns we observed may not hold true for prairie fens with greater areas than those sampled here

(i.e., 28.4 ha). Further sampling is needed in sites larger than 28.4 ha to more thoroughly examine whether sampling at 25 quadrats per site is appropriate for larger sites or if an upper and lower limited area-proportional method is more efficient.

The methods employed on this project provide acceptable levels of coverage and the ability to compare across fens. Although researchers have an understandable desire to reduce sample sizes to minimize costs and resource requirements, our data support a rigorous approach to sampling. Methodology employed in this study could be applied to future research, restoration monitoring, and conservation planning efforts being implemented in Midwestern prairie fens.

Acknowledgments The authors thank C. May and K. Pangle for their advice and expertise throughout the project and critical reading of the manuscript, and A.A. Reznicek for assistance in plant identification. We also thank all the private property owners, Jackson County, Livingston Land Conservancy, The Nature Conservancy, North Oakland Headwaters Land Conservancy, Oak Point Country Club, Servants of the Word, State of Michigan, and Washtenaw County for granting permission for research on their land. The authors also thank the editors and anonymous reviewers for their constructive comments on this manuscript. This work was supported by Botanical Society of America, Central Michigan University, Hanes Trust Foundation, Michigan Garden Club, Inc., and The Nature Conservancy. This paper is Contribution Number 70 of the Central Michigan University Institute for Great Lakes Research.

References

Albert DA (1995) Regional landscape ecosystems of Michigan, Minnesota, and Wisconsin: a working map and classification. United States Department of Agriculture, Forest Service, North Central Forest Experiment Station. Jamestown, ND: Northern Prairie Wildlife Resarch Center Online

Alroy J (2010a) The shifting balance of diversity among major marine animal groups. Science 329:1191–1194

Alroy J (2010b) Geographical, environmental and intrinsic biotic controls on Phanerozoic marine diversification. Palaeontology 53:1211–1235

Amon JP, Thompson CA, Carpenter QJ, Miner J (2002) Temperate zone fens of the glaciated Midwestern USA. Wetlands 22:301–317

Bassett T (2004) Prairie fen preserved in an urban environment. Proceedings of the 19th North American Prairie Conference 139–148

Bedford BL, Godwin KS (2003) Fens of the United States: distribution, characteristics, and scientific connection versus legal isolation. Wetlands 23:608–629

Bourdaghs M (2012) Development of a rapid floristic quality assessment. wq-bwm2-02a. Minnesota Pollution Control Agency, St Paul, pp 1–56

Bourdaghs M, Johnston CA, Regal RR (2006) Properties and performance of the floristic quality index in Great Lakes coastal wetlands. Wetlands 26:718–735

Carpenter QJ (1995) Toward a new definition of calcareous fen for Wisconsin (USA). Dissertation, University of Wisconsin-Madison

Chao A, Jost L (2012) Coverage-based rarefaction and extrapolation: standardizing samples by completeness rather than size. Ecology 93:2533–2547

Charney N, Record S (2009) Package "vegetarian." In: R Foundation for Statistical Computing. http://cran.uvigo.es/web/packages/vegetarian/vegetarian.pdf. Accessed 29 Oct 2012

Colwell RK, Coddington JA (1994) Estimating terrestrial biodiversity through extrapolation. Philos Trans R Soc B 345:101–118

Colwell RK, Mao CX, Chang J (2004) Interpolating, extrapolating, and comparing incidence-based species accumulation curves. Ecology 85:2717–2727

Colwell RK, Chao A, Gotelli NJ et al (2012) Models and estimators linking individual-based and sample-based rarefaction, extrapolation and comparison of assemblages. J Plant Ecol 5:3–21

Connor EF, McCoy ED (1979) The statistics and biology of the species-area relationship. Am Nat 113:791–833

Crancer CJ (2011) Restoration and vegetation response in Kirk Fen, a prairie fen in Ann Arbor, Michigan. Thesis, School of Natural Resources, University of Michigan

Daubenmire R (1959) A canopy-coverage method of vegetational analysis. Northwest Sci 33:43–64

Fiedler AK, Landis DA (2012) Biotic and abiotic conditions in Michigan prairie fen invaded by glossy buckthorn (*Frangula alnus*). Nat Area J 32:41–53

Francis CM, Austen M, Bowles JM, Draper WB (2000) Assessing floristic quality in southern Ontario woodlands. Nat Area J 20:66–77

Gotelli NJ, Colwell RK (2001) Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. Ecol Lett 4:379–391. doi:10.1046/j.1461-0248.2001.00230.x

Gotelli NJ, Colwell RK (2011) Estimating species richness. In: Magurran AE, McGill BJ (eds) Biological diversity: frontiers in measurement and assessment. Oxford University Press, Oxford, pp 39–54

Herman DK, Masters LA, Penskar MR et al (2001) Floristic quality assessment with wetland categories and examples of computer application for the state of Michigan, 2nd edn. Michigan Department of Natural Resources Wildlife Edition, Lansing

Hill JL, Curran PJ, Foody GM (1994) The effect of sampling on the species-area curve. Glob Ecol Biogrogr Lett 4:97–106

Horn HS (1966) Measurement of "overlap" in comparative ecological studies. Am Nat 100:419–424

Hsieh TC, Ma KH, Chao A (2013) iNEXT online: interpolation and extrapolation (version 1.3.0) [software]. http://chao.stat.nthu.edu.tw/blog/software-download/. Accessed 20 Feb 2014

Jaccard P (1912) The distribution of flora in the alpine zone. New Phytol 11:37–50
Jenks GF (1967) The data model concept in statistical mapping. Int Yearb Cartogr 7:186–190
Johnston CA, Zedler JB, Tulbure MG et al (2009) A unifying approach for evaluating the condition of wetland plant communities and identifying related stressors. Ecol Appl 19:1739–1757
Jost L (2006) Entropy and diversity. Oikos 113:363–375
Jost L (2007) Partitioning diversity into independent alpha and beta components. Ecology 88:2427–2439
Jost L (2010) The relation between evenness and diversity. Divers 2:207–232
Kost MA, Hyde DA (2009) Exploring the prairie fen wetlands of Michigan. Michigan State University, East Lansing
Kost MA, Albert DA, Cohen JG et al (2007) Natural communities of Michigan: classification and description. pp 1–323
Lopez RD, Davis CB, Fennessy MS (2002) Ecological relationships between landscape change and plant guilds in depressional wetlands. Landsc Ecol 17:43–56
Matthews JW (2003) Assessment of the floristic quality index for use in Illinois, USA, wetlands. Nat Area J 23:53–60
Matthews JW, Tessene PA, Wiesbrook SM, Zercher BW (2005) Effect of area and isolation on species richness and indices of floristic quality in Illinois, USA wetlands. Wetlands 25:607–615
Michigan Natural Features Inventory (2011) Biotics 4—Michigan's Natural Heritage Database. Lansing, Michigan. Accessed 14 Nov 2011
Minnesota Pollution Control Agency (MPCA) (2014) Rapid floristic quality assessment manual. wq-bwm2-02b. Minnesota Pollution Control Agency, St Paul, pp 1–44
Moran RC (1981) Prairie fens in northeastern Illinois: floristic composition and disturbance. In: Stuckey RL, Reese KJ (eds) Proceedings Sixth North American Prairie Conference. Columbus, OH, pp 164–168
Rasmussen SL, Starr N (1979) Optimal and adaptive stopping in the search for new species. J Am Stat Assoc 74:661
Reznicek AA, Voss EG, Walters BS (2011) Michigan Flora online. In: University of Michigan. Web. http://michiganflora.net. Accessed 16 Feb 2014
Ruiz-Jaén MC, Aide TM (2005) Restoration success: how is it being measured? Restor Ecol 13:569–577
Rydin H, Jeglum J (2006) The biology of peatlands. Oxford University Press, New York
Shannon C, Weaver W (1949) The mathematical theory of information. University of Illinois Press, Urbana
Sluis WJ (2002) Patterns of species richness and composition in re-created grassland. Restor Ecol 10:677–684
Sørensen T (1948) A method of establishing groups of equal amplitude in plant sociology based on similarity of species and its application to analyses of vegetation on Danish commons. Biol Skr Dan Vid Sel 5:1–34
Spieles JB, Comer PJ, Albert DA, Kost MA (1999) Natural community abstract for prairie fen. Michigan Natural Features Inventory, Lansing, Michigan
Swink F, Wilhelm GS (1994) Plants of the Chicago Region, 4th edn. Indiana Academy of Science, Indianapolis
Tilman D, Naeem S, Knops JMH et al (1997) Biodiversity and ecosystem properties. Science 278:1866–1867
Tilman D, Isbell F, Cowles JM (2014) Biodiversity and ecosystem functioning. Annu Rev Ecol Evol Syst 45:471–493. doi:10.1146/annurev-ecolsys-120213-091917
Tucker G, Fasham M, Hill D et al (2005) Planning a programme. In: Hill D, Fasham M, Tucker G et al (eds) Handbook of biodiversity methods: survey, evaluation and monitoring. Cambridge University Press, Cambridge, pp 6–64
Wilhelm GS, Masters LA (1995) Floristic quality assessment in the Chicago Region and application computer programs. Morton Arboretum, Lisle
Wolda H (1981) Similarity indices, sample size and diversity. Oecologia 50:296–302
Young TP (2000) Restoration ecology and conservation biology. Biol Conserv 92:73–83