ARTICLE

What is the most efficient and effective method for long-term monitoring of alpine tundra vegetation?

Steven D. Mamet, Nathan Young, Kwok P. Chun, and Jill F. Johnstone

Abstract: Nondestructive estimations of plant community characteristics are essential to vegetation monitoring programs. However, there is no universally accepted method for this purpose in the Arctic, partly because not all programs share the same logistical constraints and monitoring goals. Our aim was to determine the most efficient and effective method for long-term monitoring of alpine tundra vegetation. To achieve this, we established 12 vegetation-monitoring plots on a south-facing slope in the alpine tundra of southern Yukon Territory, Canada. Four observers assessed these plots for vascular plant species abundance employing three methods: visual cover (VC) and subplot frequency (SF) estimation and modified point-intercept (PI) (includes rare species present but not intersected by a pin). SF performed best in terms of time required per plot and sensitivity to variations in species richness. All methods were similarly poor at estimating relative abundance for rare species, but PI and VC were substantially better at high abundances. Differences among methods were larger than among observers. Our results suggest that SF is best when the monitoring focus is on rare species or species richness across extensive areas. However, when the focus is on monitoring changes in relative abundance of common species, VC or PI should be preferred.

Key words: point-intercept, subplot frequency, visual cover estimation, alpine, vegetation monitoring.

Résumé : Les estimations non destructives des caractéristiques des associations végétales sont essentielles aux programmes de surveillance de la végétation. Il n'y a cependant pas de méthode universellement acceptée à cette fin dans le cas de l'Arctique, en partie parce que tous les programmes ne partagent pas les mêmes contraintes logistiques et les mêmes objectifs de surveillance. Notre but était de déterminer la méthode la plus efficace et efficiente en ce qui a trait à la surveillance à long terme de la végétation alpine de la toundra. Dans le but de réaliser ceci, nous avons déterminé 12 parcelles de surveillance de la végétation sur un versant sud de la toundra alpine dans le sud du Territoire du Yukon, Canada. Quatre observateurs ont évalué des espèces de plantes vasculaires dans ces parcelles en employant trois méthodes : l'estimation visuelle de la couverture (VC) et de la fréquence dans la parcelle secondaire (FS) et le point d'intersection modifié (PI) (comprend les espèces rares présentes, mais non interceptées au moyen d'une tige). La méthode FS a donné les meilleurs résultats en matière du temps requis par parcelle et de sensibilité aux variations de la richesse des espèces. Toutes les méthodes étaient similairement médiocres dans l'estimation de l'abondance relative des espèces rares, mais les méthodes PI et VC étaient considérablement meilleures en matière des abondances élevées. Les différences entre les

S.D. Mamet, N. Young, and J.F. Johnstone. Northern Plant Ecology Lab, Department of Biology, University of Saskatchewan, Saskatoon, SK S7N 5E2, Canada.
Kwok P. Chun. Department of Geography, Hong Kong Baptist University, Kowloon Tong, Hong Kong.
Corresponding author: Steven D. Mamet (e-mail: sdmamet@gmail.com).

méthodes étaient plus grandes que celles entre les observateurs. Nos résultats suggèrent que la FS est meilleure lorsque l'accent de la surveillance est mis sur les espèces rares ou la richesse des espèces sur de vastes régions. Néanmoins, lorsque l'accent est mis sur la surveillance des changements de l'abondance relative des espèces communes, on devrait favoriser VC ou PI.

Mots-clés : point d'intersection, fréquence dans la parcelle secondaire, estimation visuelle de la couverture, alpin, surveillance de la végétation.

Introduction

Recent rapid temperature increases are affecting the structure and function of arctic plant communities (Sturm et al. 2001; Walker et al. 2006). However, our understanding of these effects is limited by insufficient long-term monitoring (Hudson and Henry 2010). Nondestructive estimations of plant community characteristics are essential to monitoring programs, as they allow for repeated measurements over time (Chen et al. 2009; Kopecký and Macek 2015). However, there is no universally accepted method for this purpose in the Arctic, partly because not all monitoring programs share the same logistical constraints and monitoring goals (e.g., Roberts-Pichette and Gillespie 1999; Carlson et al. 2013). For example, where time is especially limiting (e.g., the objective is sample a wide breadth of sites in a brief time period), a time-efficient method (Archaux et al. 2006) or one that requires less observer training (Archaux et al. 2009) may be desired. When monitoring for rare species, the sensitivity of a method to species richness may be more important than its ability to precisely measure abundance (Stohlgren et al. 1998), while the reverse may be true when monitoring for changes in species dominance or abundance (Vittoz and Guisan 2007). Reproducibility among observers is a critical consideration for studies of broad temporal or spatial scope but less important when designing small-scale studies (Beard et al. 1999). Since the best monitoring method may be project specific, understanding the strengths and weaknesses among methods is important for designing a successful monitoring program.

While different monitoring projects vary in their definition of "abundance" (e.g., see Kent 2011), many are interested in estimating aboveground plant biomass. Since this a destructive measure, a nondestructive proxy is needed to facilitate multiyear measurements. Point-intercept (PI), visual cover (VC), and subplot frequency (SF) are common methods of estimating plant abundance in arctic and alpine tundra (see Mueller-Dombois and Ellenberg (1974) for a description of methods). Across a broad range of ecosystem types, VC and SF are generally more time efficient than PI (Stampfli 1991; Bråkenhielm and Qinhong 1995; Prosser et al. 2003). However, some studies of low-stature vegetation in alpine (Friedmann et al. 2011) and plains (Symstad et al. 2008) environments provide evidence for the contrary, and no method has been proven ideal in all situations (Kent 2011). SF and VC may have a greater species detection rate than PI (Symstad et al. 2008; Friedmann et al. 2011), which can miss rare species simply due to random sampling error, although PI detection error can be mitigated by recording of rare species, which is the case in our study. Several studies have assessed the agreement between estimates of abundance and biomass for VC (Krebs et al. 2003; Muukkonen et al. 2006) and PI (Jonasson 1988; Frank and McNaughton 1990), but few studies have assessed this relationship for SF or compared agreement among methods. VC typically has low reproducibility among observers (Kennedy and Addison 1987; Klimeš 2003; Helm and Mead 2004) and cover estimates may not be indicative of true values (Floyd and Anderson 1987). Yet some studies have reported VC to be more reproducible than PI (Dethier et al. 1993; Symstad et al. 2008) and (or) SF (Bråkenhielm and Qinhong 1995). Efforts to improve the power and efficiency of tundra monitoring programs would thus benefit from further comparative studies that assess factors like observer bias, measurement agreement, and time requirement across multiple methods.

Faced with the challenge of identifying an optimal method of vegetation monitoring for a new field experiment (Johnstone et al. 2013), we designed a study to compare the performance of PI, VC, and SF in the context of nondestructive, long-term monitoring of alpine tundra vegetation. We destructively harvested sample plots at the end of the study to compare these nondestructive field methods against an absolute measure of abundance, namely vegetation biomass. We were particularly interested in which methods were most time-efficient and effective at estimating species composition and biomass for monitoring changes in alpine tundra vegetation. We directly compared PI, VC, and SF in terms of (*i*) field sampling time, (*ii*) sensitivity to species richness, (*iii*) agreement in estimating species abundance across different growth forms or abundance groups, and (*iv*) reproducibility among observers. In addition to informing our own monitoring decisions, our results contribute valuable information on matching measurement methods to monitoring objectives that is applicable to other research in low-stature tundra communities.

Methods

Study site and experimental design

We conducted the study in alpine tundra (1560 m asl) near Whitehorse, Yukon, Canada, at Wolf Creek (60°33.6′ N, 135°7.8′ W (Fig. 1). Previously glaciated, rounded mountain ridges in the area are overlain by fractured rock and glacial till and support tundra communities typical of midarctic shrub–heath or alpine fellfield communities (Bliss and Matveyeva 1992).

Fig. 1. Location of the study area on a south-facing slope near Whitehorse, Yukon Territory, Canada. The 12 vegetation plots investigated in this study were established immediately downslope of the vegetation-monitoring site depicted in the photograph. The wind turbine in the photograph is approximately 6 m in height. Photograph taken 14 July 2013.
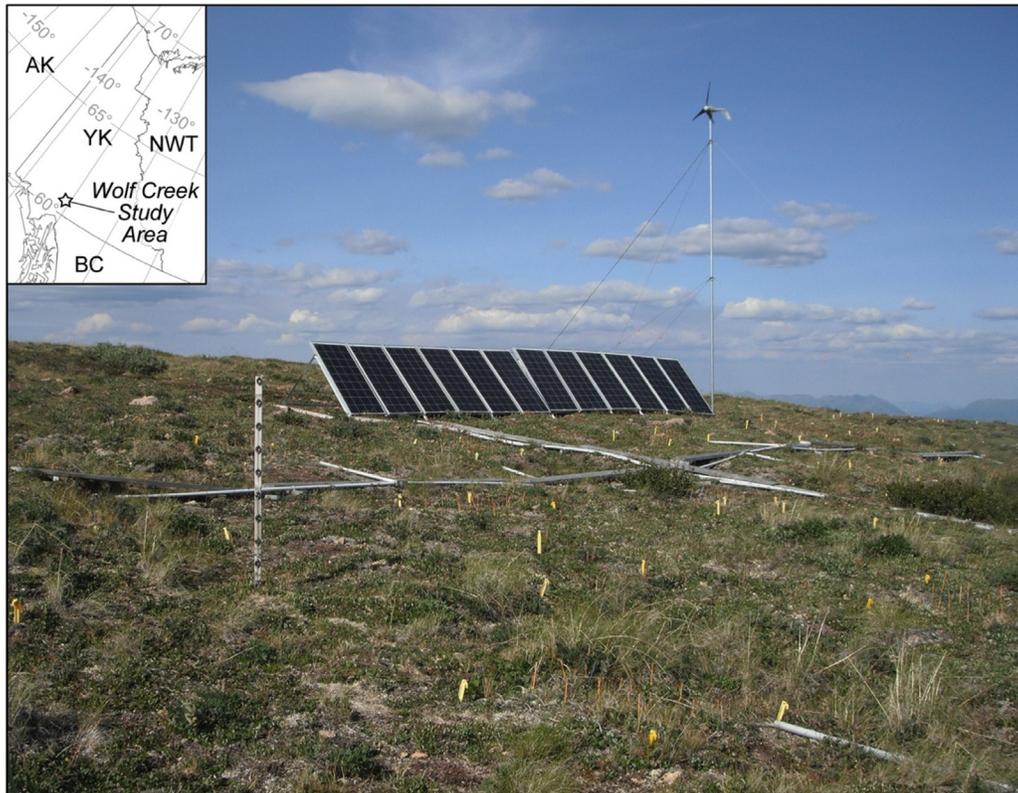
The site was dominated by low shrubs *Dryas octopetala* L., *Salix reticulata* L., and *Salix arctica* Pall., and abundant ground lichens. The vascular plant community included 35 other low-growing shrubs, forbs, grasses, and sedges that are common across North American arctic and alpine environments (Molau and Mølgaard 1996; Arft et al. 1999; Pieper et al. 2011).

Twelve vegetation plots were selected within a 4 ha area to capture the range of vegetation structure (10%–60% low-stature vascular plant cover ranging in aboveground biomass from 13.4 to 158.1 g/m^2) and dominant growth forms (forbs, graminoids, and shrubs) in low-shrub tundra near two established experimental sites (cf. Pieper et al. 2011; Johnstone et al. 2013; Fig. 1). Six plots were randomly selected along a transect running parallel to the slope contour (approximately east–west) adjacent to the experimental sites. An additional six plots were selected from the upper ($n = 3$) and lower ($n = 3$) portions of the hillslope using random compass bearings and distances from an arbitrary central point to capture the broader range of vegetation patches present across the hillslope.

Field observations

We collected vegetation data in mid-July 2011 within the period of fully expanded leaves for plants at the site. Kent (2011) recommends minimum plots sizes of 0.5 m × 0.5 m for bryophyte and lichen communities and 1 m × 1 m to 2 m × 2 m for grassland and dwarf heaths. Considering the range in plant types in our plots and maximum patch sizes (0.6 m^2), we feel 1 m × 1 m plots were an appropriate size given the nature of the vegetation at the study site. Twelve 1 m × 1 m vegetation sample plots were assessed for vascular plant species abundances by (*i*) four observers using VC and SF, (*ii*) two of the four observers using a modified PI approach (see below), and (*iii*) being destructively sampled at the end of the observation period to obtain a direct measure of plant biomass that we used as a measure of "true" plant abundance (sensu MacDonald et al. 2012). The sampling sequence occurred in a haphazard order without preconceived bias. The unbalanced sampling design between *i* and *ii* above arose due to the constraints of field logistics, as two of the four observers were needed to complete other field research tasks. We recorded the time required for plot survey for all combinations of observers, methods, and plots.

Four observers with three experience levels participated in this study. Observer 1 had extensive experience with the use of methods (>20 years of arctic plant identification experience) as well as with identification of plant species at the site. Observer 2 was a novice observer who had received detailed training in plant identification prior to the study (<1 year of plant identification experience). These two observers used all methods. Observers 3 and 4 had no explicit training with methods or plant identification until at the site and only used VC and SF. To minimize discrepancies in the analysis resulting from the incorrect identification of plant species, we provided all observers with species lists and voucher specimens of species collected previously at the site.

For SF measurements, a 1 m × 1 m quadrat was divided into 10 subsections measuring 0.5 m × 0.2 m and was placed on the surface of the plot. Species frequency was recorded as the number of subsections in which that species occurred (including at least one entire green leaf of the species); hence, SF scores ranged from 0 to 10 for each species (Kent 2011). When using this method, no community parameters other than the SF scores of vascular plant species were recorded.

Using VC, observers visually estimated the portion of ground that was covered by green tissue of a given species, as viewed from directly above the plot (Kent 2011). The abundance of lichens, mosses, litter, and bare rock was also recorded to ensure that total abundance summed to approximately 100%, but these categories were not analyzed further. Abundance was estimated to the nearest 1% (1 dm^2) for all species except those with total cover <1 dm^2, which were assigned a value of 0.1% (0.1%–0.4%) or 0.5% (0.5%–0.9%)

(Kent 2011). This approach represents a compromise between reasonably sensitive detection of rare species and moderate measurement times (Mueller-Dombois and Ellenberg 1974). We chose not to use categorical cover scales like cover abundance (Braun-Blanquet 1932), Domin–Krajina cover (Krajina 1933, cited in Mueller-Dombois and Ellenberg 1974), or canopy coverage (Daubenmire 1959) because we explicitly wished to compare observer error across a broad cover range (1% to up to 60%) using a common level of precision (to the nearest 1%).

Our PI measurements were made by recording all intersections of a pin with green vegetation (by species) as the pin was lowered to the soil surface (Molau and Mølgaard 1996). Pin intersections at 100 evenly spaced points were recorded within each plot using a gridded 1 m × 1 m frame positioned directly above the plot on four adjustable legs (Kent 2011). Pin contacts with woody, standing dead vascular plant tissue, as well as lichens, mosses, litter, and bare rock, were recorded but not included into our analyses. The greatest drawback of PI is that it can miss species, particularly rare ones (Dethier et al. 1993; Bråkenhielm and Qinhong 1995). To assist detection of rare species and intermethod comparison, we modified the PI method so that vascular plant species present in the plot but not intersected by a pin were assigned a value of one hit (Molau and Mølgaard 1996). To reduce observers' influence on each other's work, the point frame was repositioned between surveys, even when on the same plot.

Biomass is often a principle factor of interest when relating changes in vegetation to ecosystem implications (e.g., Sturm et al. 2001; Chen et al. 2009) and there are few, if any, alternatives to biomass for investigations of community structure (Chiarucci et al. 1999). Therefore, we used biomass as our reference standard of "true" plant species abundance and richness in evaluating the performance of different methods and observers in this study. Accordingly, once all of the nondestructive measurements had been completed, all aboveground biomass was removed from plots using hand clippers. The green biomass (live, nonwoody tissue) was later sorted by species in the laboratory, dried at 60°C for 48 h, and weighed to the nearest 0.001 g to capture biomass variations even at the lowest end of the abundance range.

Data analyses

To test differences among observers and methods with respect to time requirements and number of species detected in each plot (species richness), we used generalized linear mixed models in different levels of grouping factors and plot as a random effect using the R package nlme version 3.1-119 in R version 3.2.2 (Pinheiro et al. 2015; R Core Team 2016). Significant ANOVA results were followed by Tukey's honestly significant difference (HSD) test to identify pairwise differences among methods when significant main or interactive effects were detected. As only two observers performed estimates using PI, while four used VC and SF, these analyses were repeated twice: (*i*) considering only Observers 1 and 2 but all three methods and (*ii*) considering all four observers but only the SF and VC methods. For the species richness analyses, we included true richness (obtained from biomass harvest) in the methods comparison.

We used Pearson's correlation coefficient (r), estimated from linear regression, to assess the effect of method on sampling agreement, defined as the relationship between a species' estimated abundance and its actual measured biomass. We estimated sensitivity of the methods to variations in biomass as the slope of the linear regression line. Raw estimates of relative abundance (cover, pin intersections, or SF) were standardized by the maximum observed value for that method, so all methods had the same range [0,1], allowing us to compare parameter estimates between methods. Species estimates that had no corresponding biomass measure (likely due to small variations in plot positioning) were excluded from these analyses. As there were two plots where one species was not captured by the biomass

harvest $(2/(44 \text{ species} \times 12 \text{ plots}) = 0.4\%)$, we feel this decision had a negligible effect on the results. Biomass was used as the measure of actual abundance and also standardized to a proportion of the maximum observed biomass prior to analysis. This allowed direct comparison of linear regression parameters among methods.

There were obvious departures from homoscedasticity present in the data that we could not resolve using common data transformations (such as log, square root, etc.). To remedy this, we rank-transformed both the actual and estimated abundance data and compared regression parameters among methods and observers. Rank transformation is a conservative method that may alleviate problems caused by outliers and skewed data distributions that may bias results of linear regression on untransformed data (Iman and Conover 1979). As the regression results from both the rank-transformed and standardized data did not significantly differ from one another, here we only present the results of the untransformed regression analyses.

To facilitate coherent comparisons of model parameters among species or groups, we used linear models with the intercepts set at zero. The slope and correlation coefficient of the regression between actual and estimated abundance were estimated separately for each species or group (depending on the analysis), observer, and method. Species were grouped into abundance classes corresponding to a logarithmic scale of mean biomass: high (mean > 10 g/m^2), medium (1 < mean ≤10 g/m^2), low (0.1 < mean ≤1 g/m^2), and trace (mean ≤ 0.1 g/m^2). We also grouped species into general growth form types corresponding to shrubs (plants with woody stems and deciduous or evergreen leaves), graminoids (herbaceous monocots in the Poaceae, Juncaceae, or Cyperaceae family), and forbs (nongraminoid, herbaceous dicots).

To compare the reproducibility of estimating plot-level composition, we calculated Bray–Curtis (Bray and Curtis 1957) and Morisita-Horn (Horn 1966) similarity indices among observers and between observers and biomass harvest for each plot. Plant species were first grouped by growth form (i.e., shrubs, graminoids, and forbs) and then similarity indices were computed using abundance data for all species within a type. The Morisita–Horn index is more sensitive to the abundance of the most abundant species, whereas the Bray–Curtis index is more sensitive to species richness (Magurran 1988). A paired *t*-test, where each plot was a sample, was used to compare the compositional similarity of the observers among sampling methods.

To address potential issues related to small sample sizes, we employed a post hoc power analysis using the R package pwr version 1.1-3 (Champely 2015). The analysis suggested we had between a 72% and 99% chance of detecting a significant difference among methods or observers, depending on the statistical test — near the suggested benchmark of 80% (Cohen 1988). Given the relatively high power (>70%) of our statistical tests and that benchmarks should be treated with caution (Nakagawa and Cuthill 2007), we feel reasonably confident in our ability to make accurate inferences from our data set.

Results

Time requirement

The time required to estimate species abundance for the three different methods ranged from 3 to 40 min among all plots (Fig. 2). Measurements using the PI method typically required 20–30 min to complete per plot, SF took on the order of 5–10 min, and VC typically took 5–20 min. When comparing time requirements across all three methods (Observers 1 and 2 only), the effect of method ($F_{2,66} = 85.103$, $P < 0.001$) was large relative to observer effect ($F_{1,66} = 4.236$, $P = 0.044$); there was no significant interaction between method and observer ($F_{2,66} = 1.472$, $P = 0.237$). For these two observers, the mean time required to perform PI (23.4 ± 1.0 min, mean ± SE) was significantly greater than the mean time required to perform VC (8.6 ± 1.5, Tukey test: $P < 0.001$) and SF (6.3 ± 0.4, $P < 0.001$). Within a method, there were no significant differences among observers for PI or SF ($P > 0.062$).

Fig. 2. Boxplot summary of time required to complete vegetation sampling at plots in relation to observer and method. Letters below the boxplots indicate significant differences (P < 0.05) between methods as determined by generalized linear mixed models, and letters above the observer legend indicate significant differences (P < 0.05) between observers. Methods or observers that share a letter are not significantly different. Asterisks indicate significant differences within methods (P < 0.05).
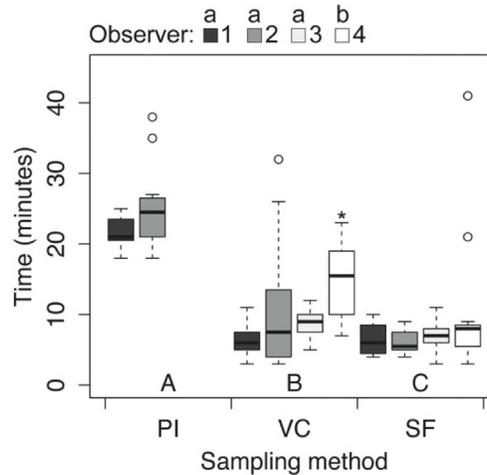
Using the VC method, Observer 4 typically took 8.2 min longer ($F_{3,44}$ = 4.433, P = 0.006) than Observer 1 to complete a plot (Fig. 2).

When considering all observers (SF and VC methods only), the main effects of method and observer were of similar strength, without a significant interaction (method, $F_{1,88}$ = 5.058, P = 0.027; observer, $F_{3,88}$ = 5.496, P = 0.002; method × observer, $F_{3,88}$ = 0.797, P = 0.499). The mean time required to perform VC was greater than the mean time for SF (10.2 ± 0.9 versus 7.6 ± 0.8, P = 0.027) (Fig. 2). Among observers, the mean time required for Observer 4 to perform estimates (12.8 min) was greater than the mean time for Observers 1 (6.5 min, P = 0.001), 2 (8.3 min, P = 0.039), and 3 (7.9 min, P = 0.018) (Fig. 2).

Species richness

Species richness determined from the biomass harvest (i.e., true richness) ranged from 9 to 18 vascular plant species/m^2 (Fig. 3). Species richness estimated by SF was comparable to true richness (P = 0.101), while PI and VC significantly underestimated species richness (P < 0.001). Richness estimates from SF and PI did not differ from each other (P = 0.187) and gave estimates that were closer to the estimates from the biomass harvest. VC gave significantly lower estimates of richness than SF (P < 0.001) but was comparable to PI (P = 0.342) (Fig. 3). Estimates of species richness did not differ significantly among observers (P > 0.050), but varied between observers 2–4 and the biomass harvest (P < 0.001). Only species estimations by Observer 1 were similar to the biomass harvest (P = 0.101).

Abundance and biomass agreement

Slopes of regression lines relating estimated relative abundance to relative biomass provided a measure of agreement (0 = no agreement, 1 = perfect agreement) and did not differ significantly among observers (P > 0.3) but did show significant differences among methods (P < 0.004). The SF method performed poorly relative to PI and VC, as indicated by significantly shallower regression slopes for SF (P = 0.003), but slopes were similar between PI and VC (P = 0.55). Similar to regression slopes, correlation coefficients relating estimated relative abundance to relative biomass indicated no significant differences among observers (P > 0.9) but significant differences among methods (P < 0.01). Correlation coefficients

Fig. 3. Boxplot summary of true species richness (i.e., biomass harvest) and species richness estimates obtained using subplot frequency (SF), point-intercept (PI), and visual cover (VC) estimates. Points represent data recorded by each observer in each plot. Letters below the boxplots indicate significant differences between methods as determined by generalized linear mixed models (P < 0.05). Methods that share a common letter are not significantly different. Richness estimates did not differ significantly among observers (P > 0.050).
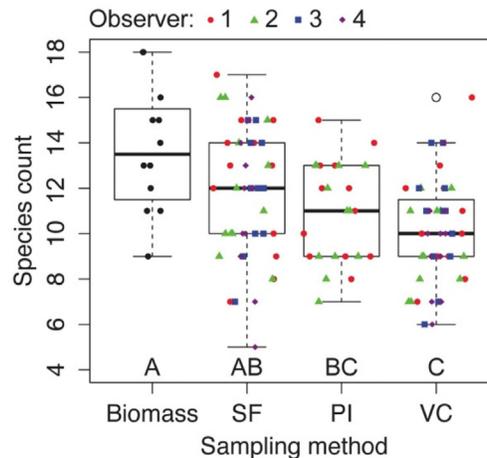
were lowest for the SF method (mean $r = 0.64$ for four observers), substantially higher for VC (mean $r = 0.95$ for four observers), and slightly higher for PI (mean $r = 0.97$ for two observers). Because we found no significant differences among observers, we concentrated our subsequent analyses on methods and plant groups.

When species were split into growth form types (forbs, graminoids, and shrubs), regression slopes for all three growth form types differed between SF and VC ($P < 0.045$) but were similar between PI and the other methods ($P > 0.063$) (Fig. 4). Correlation coefficients showed a significant interaction between growth form and method ($P = 0.026$): shrubs and forbs had similar correlation coefficients for the PI and VC methods and graminoids had lower correlation coefficients than the other growth forms for all methods (Fig. 4; Table 1).

Species abundance may also affect the relationship between a proxy estimator and actual species biomass. Analyses of regression parameters by abundance category (all methods pooled) showed differences in both slope ($P < 0.001$) and correlation coefficient ($P < 0.001$). Regression slopes for abundant species were significantly steeper than for rare species (Fig. 5). Correlation coefficients also showed a significant interaction between abundance class and method ($P < 0.001$), with unique patterns of correlation coefficients across categories for each method (Fig. 5). SF estimates of plant abundance consistently showed the least agreement with biomass, with the exception of trace species in which all methods performed poorly (i.e., shallow slopes).

Reproducibility

Forb and shrub growth forms, when observed using PI and VC methods, showed significant and high multivariate correlations between observers ($r \geq 0.96$, $P < 0.001$); graminoids were noticeably lower ($r \geq 0.82$, $P < 0.001$) (Table 2). In contrast, the SF method yielded lower multivariate correlations ($r < 0.70$) although it was better for shrubs ($r \geq 0.77$).

Discussion and conclusion

Time requirement

Our results suggest that the three methods of measuring plant abundance differ substantially in the amount of time required to survey an individual 1 m × 1 m plot. SF and

Fig. 4. Relationships between estimated relative abundance (proportion of maximum observed value) and relative biomass (proportion of maximum observed value) across different plant growth form types. Each point represents a species observation within a plot for one of four observers; the shape of the point indicates (*a–c*) method (SF = subplot frequency, PI = point-intercept, and VC = visual cover) or (*d–f*) observer. Regression lines and slope coefficients (β) are shown separately for different observers and methods. The solid black line represents a 1:1 relationship. *P < 0.05, **P < 0.01, and ***P < 0.001 significance. Data were rank-transformed prior to analysis.
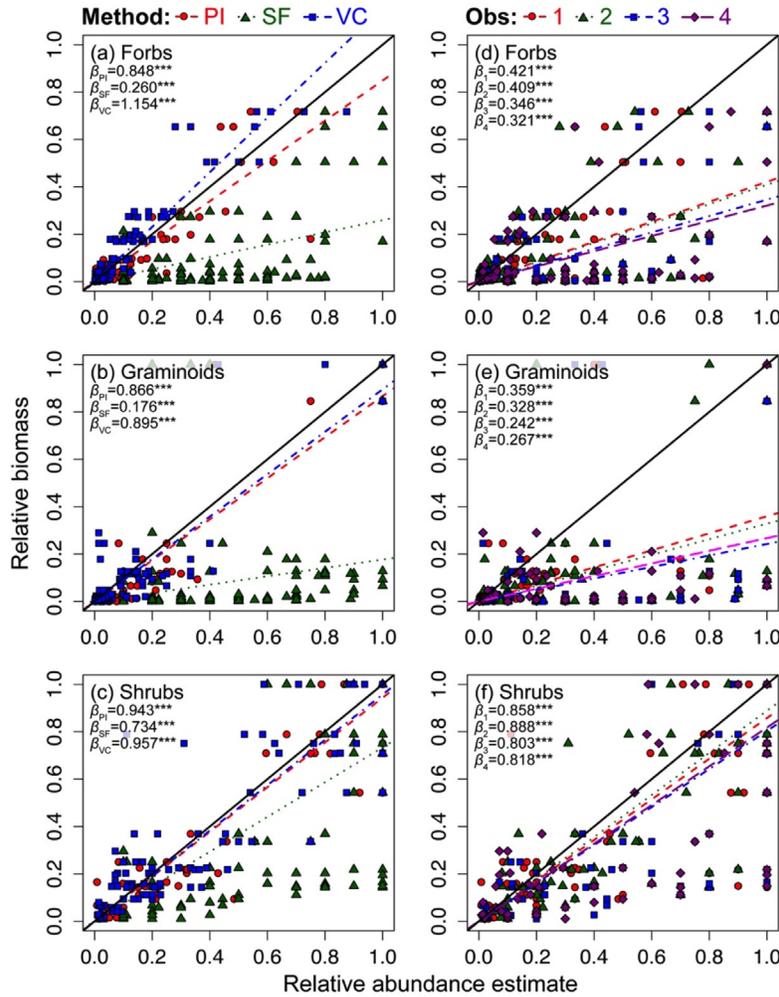

Table 1. Mean correlation coefficients for regressions of estimated relative plant abundance versus measured biomass of individual species grouped by measurement method (SF = subplot frequency, PI = modified point-intercept, VC = visual cover) and growth form (left columns) or cover class (right columns).

Method	Forbs	Graminoids	Shrubs	High	Medium	Low	Trace
PI	0.96 (0.00)	0.82 (0.02)	0.96 (0.01)	0.95 (0.01)	0.94 (0.01)	0.84 (0.02)	0.41 (0.01)
VC	0.97 (0.01)	0.82 (0.02)	0.93 (0.01)	0.90 (0.01)	0.91 (0.04)	0.81 (0.04)	0.69 (0.04)
SF	0.61 (0.01)	0.54 (0.03)	0.69 (0.02)	0.61 (0.01)	0.52 (0.02)	0.30 (0.01)	0.47 (0.02)

Note: Values are means across species and observers presented with standard errors in parentheses.

Fig. 5. Relationships between estimated relative abundance (proportion of maximum observed value, x-axis) and relative biomass (proportion of maximum observed value, y-axis) across different abundance classes based on mean plot-level species biomass (high: mean > 10 g/m^2, medium: 1 < mean < 10 g/m^2, low: 0.1 < mean < 1 g/m^2, trace: mean < 0.1 g/m^2). Each point represents a species observation within a plot for one of four observers; the shape of the point indicates (*a–d*) method (SF = subplot frequency, PI = point-intercept, and VC = visual cover) or (*e–h*) observer. Regression lines and slope coefficients (β) are shown separately for different observers and methods. The solid black line represents a 1:1 relationship. *P < 0.05, **P < 0.01, and ***P < 0.001 significance. Data were rank-transformed prior to analysis.
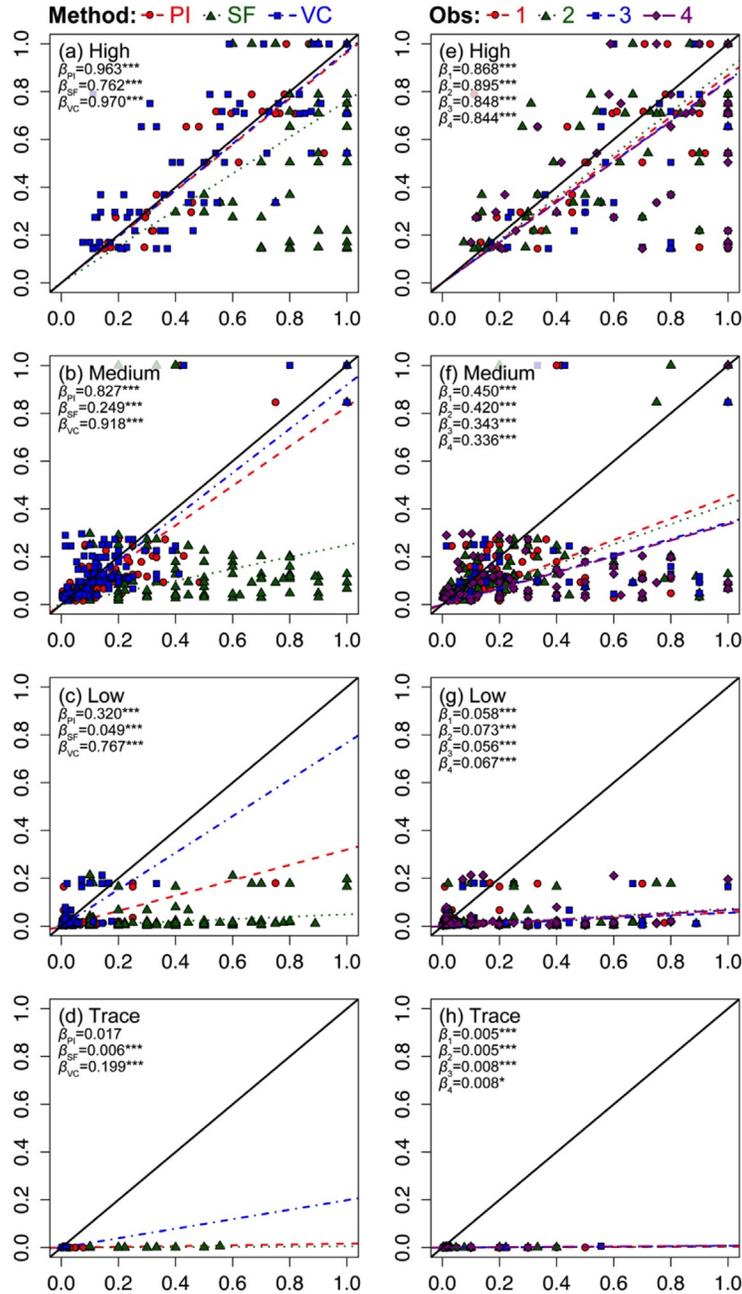
Table 2. Correlation and similarity measures between vegetation abundance measures recorded by four observers using the modified point-intercept (PI), visual cover (VC), and subplot frequency (SF) methods.

Measure		PI				VC				SF			
		Obs1	Obs2	Obs3	Obs4	Obs1	Obs2	Obs3	Obs4	Obs1	Obs2	Obs3	Obs4
Corr[a]	1	0.96	0.98			0.94	0.96	0.95	0.96	0.65	0.66	0.66	0.64
	2	0.96	0.96			0.97	0.98	0.96	0.99	0.63	0.67	0.64	0.62
	3	0.97	0.98			0.94	0.96	0.95	0.95	0.80	0.82	0.77	0.78
	4	0.87	0.82			0.88	0.82	0.84	0.83	0.60	0.67	0.65	0.60
	n[b]	131	122			129	110	124	111	142	137	138	131

		Obs:Obs		Obs:Bio		Obs:Obs		Obs:Bio		Obs:Obs		Obs:Bio	
BC[c]	1[d]	0.84 (0.02)		0.83 (0.01)		0.81 (0.01)		0.87 (0.01)		0.89 (0.01)		0.53 (0.01)	
	2	0.76 (0.03)		0.84 (0.02)		0.79 (0.02)		0.89 (0.01)		0.89 (0.01)		0.58 (0.02)	
	3	0.91 (0.01)		0.91 (0.01)		0.83 (0.02)		0.90 (0.01)		0.93 (0.01)		0.75 (0.02)	
	4	0.67 (0.05)		0.67 (0.04)		0.62 (0.03)		0.77 (0.02)		0.82 (0.01)		0.67 (0.02)	
MH[c]	1	0.97 (0.02)		0.96 (0.01)		0.96 (0.01)		0.96 (0.01)		0.96 (0.00)		0.64 (0.01)	
	2	0.92 (0.03)		0.95 (0.02)		0.96 (0.01)		0.97 (0.01)		0.96 (0.01)		0.72 (0.02)	
	3	0.99 (0.00)		0.95 (0.02)		0.95 (0.02)		0.97 (0.01)		0.98 (0.01)		0.72 (0.02)	
	4	0.80 (0.05)		0.79 (0.05)		0.85 (0.02)		0.88 (0.02)		0.93 (0.01)		0.80 (0.02)	

Note: Mean similarity values calculated among observers (Obs:Obs) and between observers and measured biomass (Obs:Bio) using multivariate abundances are presented with standard errors in parentheses.

[a]Correlation between plant abundance estimated by each observer (1–4) and biomass (g dry mass/m^2) for each plant growth type such that each point in the correlation analysis represents a unique combination of species and plot. All correlations are significant at $P < 0.001$.

[b]Total sample size (number of species × plot observations) for each observer and total abundance comparison (e.g., Obs1:Total abundance = 18 species abundances observed across 12 plots = 131 observations).

[c]BC = Bray–Curtis, MH = Morisita–Horn.

[d]1 = Total abundance, 2 = forb abundance, 3 = shrub abundance, 4 = graminoid abundance.

VC required less than half the time to implement compared to PI. In contrast, the amount of time required to complete estimates was only weakly related to differences in observer training. For example, we observed only minor differences in time efficiency between Observer 1, with extensive research experience, and Observer 2, who had training in plant identification but not methods prior to the study. We also found little difference between "experienced" (Observers 1 and 2) and "inexperienced" observers (Observers 3 and 4) in time requirements. Although based on a small sample size of four individuals, our results suggest that variations in prior experience may not translate into large differences in sampling efficiency, provided all observers receive similar training in the field.

Species richness

Unlike time requirements, estimates of species richness were more strongly influenced by observer, suggesting that observer biases are influential when monitoring for biodiversity surveys or rare species. We also found that method type affected species detection, with a significantly higher detection rate using SF and PI relative to VC across all observers. This result contrasts with other studies that have found lower species detection using PI compared to VC methods (Symstad et al. 2008; Godínez-Alvarez et al. 2009). Theoretically, PI has a tendency to miss rare species simply because of sampling error; typically, a species must be contacted by at least one pin to be registered as present in the plot, and the likelihood of this occurring decreases with the cover of a species (Friedmann et al. 2011). Our modified PI method may have partially overcome this difficulty by recording species that were present but not hit by a pin (cf. Dethier et al. 1993; Molau and Mølgaard 1996), thus capturing more rare species. We suspect that the rapidity of VC (mean = 8.6 min per plot) may have resulted in increased detection errors compared to the slower PI method (mean = 23.4 min per plot), simply because observers spent less time closely looking at each plot. VC is

also more subjective than PI, and user fatigue (e.g., hasty plant identification and (or) quantification) or lack of experience could hamper its usefulness. For example, we found that our "inexperienced" observers (Observers 3 and 4) also had the lowest species detection rate. Thus, careful observation, as well as adequate plant identification training, is crucial for maximizing the effectiveness of VC if it is intended for estimating species richness. Overall, our results suggest SF may represent the best method when the primary research objectives are tracking changes in rare species or species richness due to the high sensitivity of SF to rare species, its low time requirements, and the high reproducibility of species counts among observers.

Agreement and reproducibility of abundance estimates

Method of assessment had a significant effect on the agreement of abundance estimates both when species were pooled together and when separated by growth form. In contrast, observer variation had little effect, indicating that observer bias may be relatively minor compared to method choice. Across all observers, estimates of abundance using SF showed the weakest relationships with actual biomass. Thus, this method seems poorly suited to estimating relative abundance (cf. Bråkenhielm and Qinhong 1995). We expect that increasing the number of cells within a plot that are assessed for species frequency would increase agreement, and it is likely that this method would eventually converge to a level of precision similar to PI. However, frequency based on subplots may strongly differ from real abundance for some species — very small species may have high frequency but low abundance or species with high abundance may have low frequency (Kent 2011). Additionally, a dramatic increase in the number of sampled subcells with the SF method would likely counteract other advantages such as low sampling time. Thus, it seems reasonable to suggest that when the study aims are to estimate relative abundance of common and rare species, the SF method is not recommended, as it does not accurately estimate plant abundance or biomass.

All methods showed a trend towards greater agreement in the more abundant plant groups. Importantly, our results show that correlations with actual biomass are lower for all methods when estimating plants in the trace- (<0.1 g) and low-biomass (0.1–1 g) groups, which suggests that none of these methods are precise tools for detecting changes in abundance of very rare species. Regardless of the method used, low-abundance species are problematic because individuals are inherently difficult to detect (Klimeš et al. 2001; Klimeš 2003; Friedmann et al. 2011;). Moreover, our study did not assess optimal plot size or number of plots among methods. Although we were able to detect statistical differences among methods and observers with reasonable confidence (>70%), we have yet to determine the optimal number of plots for tundra vegetation monitoring. Future work should address these matters of plot size, frequency, and rare species to further develop effective monitoring methods of arctic and alpine tundra vegetation.

Within abundance groups, there was general concurrence between abundance estimates derived from both PI and VC. The evaluation of the best sampling approach for abundance estimation should be taken within the context of time required for sampling, as PI sampling in this study took approximately twice the time of VC sampling (Table 3). One important consideration that would affect the time requirements when using PI is the number of pin contacts used at each plot. Theoretically, using more pins increases the precision of estimates, but beyond a threshold of pins, the gains in precision associated with increasing the number of pins may be outweighed by increases in time requirement (e.g., Bråthen and Hagberg 2004; Godínez-Alvarez et al. 2009). Moreover, counting more than one hit per pin in PI estimation may overestimate relative abundances (particularly for graminoids), although more accurately estimate biomass. If the objective is to monitor species with high abundance (i.e., >50%), 100 pins/m^2 is likely sufficient, although if rare species are

Table 3. Summary of results and advantages and disadvantages of three methods of vegetation monitoring: subplot frequency (SF), point-intercept (PI), and visual cover (VC) estimation.

Method	Efficiency	Species detection	Relative abundance	Advantages	Disadvantages
SF	(1) 6.3 min	(1) 11.9 sp.	(3) $r = 0.491$***	Quick and easy to implement; high species detection	Poor quantification of low-cover species; low reproducibility
PI	(3) 23.4 min	(2) 11.0 sp.	(1) $r = 0.970$***	High species detection; high reproducibility	Time consuming; requires modification to detect rare species
VC	(2) 8.6 min	(3) 10.1 sp.	(2) $r = 0.976$***	Accurate for high-cover species; quick and easy to implement	Poor detection of rare species; low reproducibility

Note: Mean values are presented for each method with relative ranking of each method for efficiency, species detection, and relative abundance in parentheses. ***$P < 0.001$ significance.

desired, the number of pins per square metre should be increased (Dethier et al. 1993) while remaining mindful that there may be minimum gains in accuracy above 10 pins/0.25 m^2 (Bråthen and Hagberg 2004). Overall, since point techniques require very large samples for detecting change in less frequent species (either through many plots or many points per plot), we echo the recommendation of Everson et al. (1990) that key species with high abundance be selected for monitoring changes in species composition.

Some plant growth forms present additional challenges for abundance estimation. In our study, graminoids showed a unique relationship between estimated and actual biomass, indicating a need to carefully calibrate this relationship when working with this growth form. In addition, the agreement of abundance estimates tended to be lower for graminoids and highly sensitive to the type of method. The upright and highly dissected arrangement of graminoid leaves make it difficult to visually estimate the area covered during VC and may also affect pin intersection probability during PI, especially under windy conditions (Vittoz and Guisan 2007; Godínez-Alvarez et al. 2009). Multivariate indices of similarity between observer measurements and relative biomass suggest that the VC method may be better than PI or SF for estimating variations in graminoid biomass.

Conclusions

Developing a strong understanding of the strengths and weaknesses of different methods for monitoring vegetation in alpine tundra is an essential aspect of designing a successful monitoring program. We identified statistically significant relationships among relative abundance and richness estimates from three different methods (SF, PI, and VC) compared to "true" values as determined by biomass harvesting, in spite of the small number of plots sampled ($n = 12$). Our comparison of methods here suggests that SF is the best approach when monitoring objectives are focused on rare species detection or assessing species richness with its corresponding requirements for extensive sampling. However, when the focus is to monitor changes in the abundance of common species, VC or PI should be preferred. In this study, VC required less time to sample than PI, although VC sensitivity to species richness was lower than for PI. In general, differences among observers were small compared to differences among methods, indicating the choice of monitoring method should supersede potential concerns about observer bias.

Acknowledgements

References

Archaux, F., Gosselin, F., Bergès, L., and Chevalier, R. 2006. Effects of sampling time, species richness and observer on the exhaustiveness of plant censuses. J. Veg. Sci. **17**(3): 299–306. doi: 10.1111/j.1654-1103.2006.tb02449.x.

Archaux, F., Camaret, S., Dupouey, J.-L., Ulrich, E., Corcket, E., Bourjot, L., Brêthes, A., Chevalier, R., Dobremez, J.-F., Dumas, Y., Dumé, G., Forêt, M., Forgeard, F., Lebret Gallet, M., Picard, J.-F., Richard, F., Savoie, J.-M., Seytre, L., Timbal, J., and Touffet, J. 2009. Can we reliably estimate species richness with large plots? An assessment through calibration training. Plant Ecol. **203**(2): 303–315. doi: 10.1007/s11258-008-9551-6.

Arft, A.M., Walker, M.D., Gurevitch, J., Alatalo, J.M., Bret-Harte, M.S., Dale, M., Diemer, M., Gugerli, F., Henry, G.H.R., Jones, M.H., Hollister, R.D., Jónsdóttir, I. S., Laine, K., Lévesque, E., Marion, G.M., Molau, U., Mølgaard, P., Nordenhäll, U., Raszhivin, V., Robinson, C.H., Starr, G., Stenström, A., Stenström, M., Totland, Ø., Turner, P.L., Walker, L.J., Webber, P.J., Welker, J.M., and Wookey, P.A. 1999. Responses of tundra plants to experimental warming: meta-analysis of the International Tundra Experiment. Ecol. Monogr. **69**(4): 491–511. doi: 10.2307/2657227.

Beard, G.R., Scott, W.A., and Adamson, J.K. 1999. The value of consistent methodology in long-term environmental monitoring. Environ. Monit. Assess. **54**(3): 239–258. doi: 10.1023/a:1005917929050.

Bliss, L.C., and Matveyeva, N.V. 1992. Circumpolar arctic vegetation. *In* Arctic ecosystems in a changing climate: an ecophysiological perspective. *Edited by* F.S. Chapin III, R.L. Jeffries, J.F. Reynolds, G.R. Shaver, and J. Svoboda. Academic Press, San Diego, Calif. pp. 59–89.

Bråkenhielm, S., and Qinhong, L. 1995. Comparison of field methods in vegetation monitoring. Water Air Soil Pollut. **79**: 75–87. doi: 10.1007/BF01100431.

Bråthen, K.A., and Hagberg, O. 2004. More efficient estimation of plant biomass. J. Veg. Sci. **15**(5): 653–660. doi: 10.1111/j.1654-1103.2004.tb02307.x.

Braun-Blanquet, J. 1932. Plant Sociology. The study of plant communities. McGraw-Hill, New York.

Bray, J.R., and Curtis, J.T. 1957. An ordination of the upland forest communities of southern Wisconsin. Ecol. Monogr. **27**: 325–349. doi: 10.2307/1942268.

Carlson, B.Z., Randin, C.F., Boulangeat, I., Lavergne, S., Thuiller, W., and Choler, P. 2013. Working toward integrated models of alpine plant distribution. Alp Bot. **123**(2): 41–53. doi: 10.1007/s00035-013-0117-4.

Champely, S. 2015. pwr: basic functions for power analysis. R package version 1.1-3. Available from: http://CRAN.R-project.org/package=pwr

Chen, W., Li, J., Zhang, Y., Zhou, F., Koehler, K., LeBlanc, S., Fraser, R., Olthof, I., Zhang, Y., and Wang, J. 2009. Relating biomass and leaf area index to non-destructive measurements in order to monitor changes in Arctic vegetation. Arctic. **62**(3): 281–294. doi: 10.14430/arctic148.

Chiarucci, A., Wilson, J.B., Anderson, B.J., and Dominicis, V.D. 1999. Cover versus biomass as an estimate of species abundance: does it make a difference to the conclusions? J. Veg. Sci. **10**(1): 35–42. doi: 10.2307/3237158.

Cohen, J. 1988. Statistical power analysis for the behavioral sciences. 2nd ed. Lawrence Erlbaum, Hillsdale, N.J.

Daubenmire, R. 1959. A canopy-coverage method of vegetational analysis. Northwest Sci. **33**(1): 33–51.

Dethier, M.N., Graham, E.S., Cohen, S., and Tear, L.M. 1993. Visual versus random-point percent cover estimations: 'objective' is not always better. Mar. Ecol. Progr. Ser. **96**: 93–100. doi: 10.3354/meps096093.

Everson, T.M., Clarke, G.P.Y., and Everson, C.S. 1990. Precision in monitoring plant species composition in montane grasslands. Vegetatio. **88**(2): 135–141. doi: 10.1007/BF00044830.

Floyd, D.A., and Anderson, J.E. 1987. A comparison of three methods for estimating plant cover. J. Ecol. **75**(1): 221–228. doi: 10.2307/2260547.

Frank, D.A., and McNaughton, S.J. 1990. Aboveground biomass estimation with the canopy tntercept method: a plant growth form caveat. Oikos. **57**(1): 57–60. doi: 10.2307/3565736.

Friedmann, B., Pauli, H., Gottfried, M., and Grabherr, G. 2011. Suitability of methods for recording species numbers and cover in alpine long-term vegetation monitoring. Phytocoenologia. **41**(2): 143–149. doi: 10.1127/0340-269X/2011/0041-0480.

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. Ecol. Indicat. **9**(5): 1001–1008. doi: 10.1016/j.ecolind.2008.11.011.

Helm, D.J., and Mead, B.R. 2004. Reproducibility of vegetation cover estimates in south-central Alaska forests. J. Veg. Sci. **15**(1): 33–40. doi: 10.1111/j.1654-1103.2004.tb02234.x.

Horn, H. 1966. Measurement of "overlap" in comparative ecological studies. Am. Nat. **100**: 419–424. doi: 10.1086/282436.

Hudson, J.M.G., and Henry, G.H.R. 2010. High Arctic plant community resists 15 years of experimental warming. J. Ecol. **98**(5): 1035–1041. doi: 10.1111/j.1365-2745.2010.01690.x.

Iman, R.L., and Conover, W.J. 1979. The use of the rank transform in regression. Technometrics. **21**(4): 499–509. doi: 10.1080/00401706.1979.10489820.

Johnstone, J.F., Henkelman, J., Allen, K., Helgason, W., and Bedard-Haughn, A. 2013. Controlled soil warming powered by alternative energy for remote field sites. PLoS One. **8**(12): e82903. doi: 10.1371/journal.pone.0082903.

Jonasson, S. 1988. Evaluation of the point intercept method for the estimation of plant biomass. Oikos. **52**(1): 101–106. doi: 10.2307/3565988.

Kennedy, K.A., and Addison, P.A. 1987. Some considerations for the use of visual estimates of plant cover in biomonitoring. J. Ecol. **75**(1): 151–157. doi: 10.2307/2260541.

Kent, M. 2011. Vegetation description and data analysis: a practical approach. 2nd ed. Wiley, Hoboken, N.J.

Klimeš, L. 2003. Scale-dependent variation in visual estimates of grassland plant cover. J. Veg. Sci. **14**(6): 815–821. doi: 10.1111/j.1654-1103.2003.tb02214.x.

Klimeš, L., Dančak, M., Hájek, M., Jongepierová, I., and Kučera, T. 2001. Scale-dependent biases in species counts in a grassland. J. Veg. Sci. **12**(5): 699–704. doi: 10.2307/3236910.

Kopecký, M., and Macek, M. 2015. Vegetation resurvey is robust to plot location uncertainty. Divers. Distrib. **21**(3): 322–330. doi: 10.1111/ddi.12299.

Krajina, V.J. 1933. Die Pflanzengesellshaften des Mlynica-Tales in den Vysoke Tatry (HochTatra). Mit besonderer Berucksichitigung der oekologischen Verhaltnisse. Beihefte zum Botanischen Centralblatt. **2**(50–51): 744–957, 741–244

Krebs, C.J., Danell, K., Angerbjörn, A., Agrell, J., Berteaux, D., Bråthen, K.A., Danell, Ö., Erlinge, S., Fedorov, V., Fredga, K., Hjältén, J., Högstedt, G., Jónsdóttir, I.S., Kenney, A.J., Nordin, N.K., Roininen, H., Svensson, M., Tannerfeldt, M., and Wiklund, C. 2003. Terrestrial trophic dynamics in the Canadian Arctic. Can. J. Zool. **81**: 827–843. doi: 10.1139/z03-061.

MacDonald, R.L., Burke, J.M., Chen, H.Y.H., and Prepas, E.E. 2012. Relationship between aboveground biomass and percent cover of ground vegetation in Canadian Boreal Plain riparian forests. For. Sci. **58**(1): 47–53. doi: 10.5849/forsci.10-129.

Magurran, A.E. 1988. Ecological diversity and its measurement. Princeton University Press, Princeton, N.J.

Molau, U., and Mølgaard, P. 1996. International Tundra Experiment − ITEX Manual. Danish Polar Center, Copenhagen.

Mueller-Dombois, D., and Ellenberg, H. 1974. Aims and methods in vegetation ecology. Wiley, New York.

Muukkonen, P., Mäkipää, R., Laiho, R., Minkkinen, K., Vasander, H., and Finér, L. 2006. Relationship between biomass and percentage cover in understorey vegetation of boreal coniferous forests. Silva Fenn. **40**(2): 231–245. doi: 10.14214/sf.340.

Nakagawa, S., and Cuthill, I.C. 2007. Effect size, confidence interval and statistical significance: a practical guide for biologists. Biol. Rev. **82**(4): 591–605. doi: 10.1111/j.1469-185X.2007.00027.x.

Pieper, S.J., Loewen, V., Gill, M., and Johnstone, J.F. 2011. Plant responses to natural and experimental variations in temperature in alpine tundra, southern Yukon, Canada. Arct. Antarct.Alp. Res. **43**(3): 442–456. doi: 10.1657/1938-4246-43.3.442.

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., and R Core Team. 2015. nlme: linear and nonlinear mixed effects models. R package version 3.1-119. Available from: http://CRAN.R-project.org/package=nlme

Prosser, C.W., Skinner, K.M., and Sedivec, K.K. 2003. Comparison of 2 techniques for monitoring vegetation on military lands. J. Range Manage. **56**(5): 446–454. doi: 10.2307/4003835.

R Core Team. 2016. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available from: http://www.r-project.org/

Roberts-Pichette, P., and Gillespie, L.J. 1999. Terrestrial vegetation biodiversity monitoring protocols. EMAN occasional paper series. Report No. 9. Ecological Monitoring Coordinating Office, Canada Centre for Inland Waters, Burlington, Ont.

Stampfli, A. 1991. Accurate determination of vegetational change in meadows by successive point quadrat analysis. Vegetatio. **96**(2): 185–194. doi: 10.1007/BF00044979.

Stohlgren, T.J., Bull, K.A., and Otsuki, Y. 1998. Comparison of rangeland vegetation sampling techniques in the Central Grasslands. J. Range Manage. **51**(2): 164–172. doi: 10.2307/4003202.

Sturm, M., Racine, C., and Tape, K. 2001. Climate change − increasing shrub abundance in the Arctic. Nature. **411**(6837): 546–547. doi: 10.1038/35079180.

Symstad, A.J., Wienk, C.L., and Thorstenson, A.D. 2008. Precision, repeatability, and efficiency of two canopy-cover estimate methods in northern Great Plains vegetation. Rangeland Ecol. Manage. **61**(4): 419–429. doi: 10.2111/08-010.1.

Vittoz, P., and Guisan, A. 2007. How reliable is the monitoring of permanent vegetation plots? A test with multiple observers. J. Veg. Sci. **18**(3): 413–422. doi: 10.1111/j.1654-1103.2007.tb02553.x.

Walker, M.D., Wahren, C.H., Hollister, R.D., Henry, G.H.R., Ahlquist, L.E., Alatalo, J.M., Bret-Harte, M.S., Calef, M.P., Callaghan, T.V., Carroll, A.B., Epstein, H.E., Jónsdóttir, I.S., Klein, J.A., Magnússon, B., Molau, U., Oberbauer, S.F., Rewa, S.P., Robinson, C.H., Shaver, G.R., Suding, K.N., Thompson, C.C., Tolvanen, A., Totland, Ø., Turner, P.L., Tweedie, C.E., Webber, P.J., and Wookey, P.A. 2006. Plant community responses to experimental warming across the tundra biome. Proc. Natl. Acad. Sci. U.S.A. **103**(5): 1342–1346. doi: 10.1073/pnas.0503198103.