Biological Conservation 144 (2011) 1270–1277

Contents lists available at ScienceDirect

Biological Conservation

journal homepage: www.elsevier.com/locate/biocon

Developing effective sampling designs for monitoring natural resources in Alaskan national parks: An example using simulations and vegetation data

William L. Thompson [a,*], Amy E. Miller [a], Dorothy C. Mortenson [a,b], Andrea Woodward [c]

[a] National Park Service, Southwest Alaska Network, 240 West 5th Avenue, Anchorage, AK 99501, United States
[b] Oregon Water Resources Department, 725 Summer Street NE, Suite A, Salem, OR 97301, United States
[c] US Geological Survey, Forest and Rangeland Ecosystem Science Center, 6505 NE 65th Street, Seattle, WA 98115, United States

ARTICLE INFO

Article history:
Received 6 April 2010
Received in revised form 23 September 2010
Accepted 27 September 2010
Available online 27 October 2010

Keywords:
Sample size estimation
Sampling frequency
Coefficient of variation
Path Distance analysis

ABSTRACT

Monitoring natural resources in Alaskan national parks is challenging because of their remoteness, limited accessibility, and high sampling costs. We describe an iterative, three-phased process for developing sampling designs based on our efforts to establish a vegetation monitoring program in southwest Alaska. In the first phase, we defined a sampling frame based on land ownership and specific vegetated habitats within the park boundaries and used Path Distance analysis tools to create a GIS layer that delineated portions of each park that could be feasibly accessed for ground sampling. In the second phase, we used simulations based on landcover maps to identify size and configuration of the ground sampling units (single plots or grids of plots) and to refine areas to be potentially sampled. In the third phase, we used a second set of simulations to estimate sample size and sampling frequency required to have a reasonable chance of detecting a minimum trend in vegetation cover for a specified time period and level of statistical confidence. Results of the first set of simulations indicated that a spatially balanced random sample of single plots from the most common landcover types yielded the most efficient sampling scheme. Results of the second set of simulations were compared with field data and indicated that we should be able to detect at least a 25% change in vegetation attributes over 31 years by sampling 8 or more plots per year every five years in focal landcover types. This approach would be especially useful in situations where ground sampling is restricted by access.

Published by Elsevier Ltd.

1. Introduction

The United States National Park Service (NPS) is charged with protecting or maintaining natural systems within US national parks. As part of this mandate, the NPS Inventory and Monitoring (I&M) Program was developed to collect scientifically sound data on the current status and long-term trends of key components of park ecosystems. Monitoring data, in particular, are expected to promote a better understanding of ecosystem processes, and to provide information used in decision making for resource protection and management, as well as public education and outreach (Fancy et al., 2009). Monitoring data should also serve to document alterations to park ecosystems as a result of climate change and other factors (Lindenmayer and Likens, 2009).

Despite the importance of credible data on the status and trends of key resources, far too many natural resource monitoring programs to date have been ineffective because of inadequate programmatic support, poor planning, and lack of rigorous study design (Legg and Nagy, 2006; Field et al., 2007; Lindenmayer and Likens, 2009). Challenges to developing a rigorous design include the need to maximize spatial balance of a random sample of plots while minimizing sampling effort, the need for sample sizes sufficient to detect change in heterogeneous or highly variable environments, and the long time periods required to detect change. The parks in southwest Alaska are large (0.3–1.7 million ha, each), remote, and generally lack infrastructure for travel (roads, trails). Additional logistical challenges associated with sampling include a limited number of landing areas for planes to access sample areas, inclement weather, and a short summer field season.

Practical considerations, such as accessibility and cost, necessarily limit the sampling design alternatives that can be implemented in remote, roadless parks. Possible sampling approaches include remote-sensing techniques (e.g., change detection using mid- and high-resolution imagery), aerial surveys (e.g., population counts) and/or ground-based sampling (e.g., plot-based sampling). Each approach has technological, budgetary, and/or logistical constraints and addresses monitoring objectives at different temporal and spatial scales. Although remote-sensing data can provide wall-to-wall coverage of park resources at a much lower cost than field

* Corresponding author. Tel.: +1 907 644 3685; fax: +1 907 644 3809.
E-mail addresses: bill_thompson@nps.gov (W.L. Thompson), amy_e_miller@nps.gov (A.E. Miller), mortendc@wrd.state.or.us (D.C. Mortenson), andrea_woodward@usgs.gov (A. Woodward).

0006-3207/$ - see front matter Published by Elsevier Ltd.
doi:10.1016/j.biocon.2010.09.032

visits, there are many aspects of resource condition that are too subtle to detect spectrally, or occur at too fine a scale to monitor through remote-sensing techniques alone.

In this paper, we outline an iterative, three-phase process that we used to develop a sampling design, adapted from Woodward et al. (2009), for ground-based monitoring of vegetation in national parks of the Southwest Alaska Network (SWAN; Fig. 1), one of 32 I&M networks nationwide. In Alaska and other northern latitudes, regional warming is expected to contribute to changes in species composition (e.g., Walker et al., 2006; Hudson and Henry, 2009); shifts in species' ranges (e.g., Sturm et al., 2001; Lesica and McCune, 2004); drought stress and reduced tree growth (e.g., Barber et al., 2000); and increasingly severe, large-scale disturbances (e.g., Berg et al., 2006). Our overarching objective was to estimate long-term trends in the composition and structure of late-successional plant communities from three elevation bands (0–450 m; 451–900 m; >900 m) in the three largest SWAN parks. We framed our monitoring objective as a function of elevation because we expected both drivers (environmental) and response variables (vegetation attributes) to vary along an altitudinal gradient. We further separated estimates of vegetation change by park because regional effects (e.g., geography) were also expected to influence variation in vegetation composition and structure.

Our sampling design is intended to capture changes at the level of individual species or plant communities, including changes in species richness and species turnover that could be associated with climate change. We address the challenges outlined above (access; sampling effort; ability to detect change) with a focus on the methods we used to delineate our population of interest, determine sample unit configuration, and estimate sample size and frequency required to detect a change or trend. The sampling design is thus tailored to meet the monitoring objectives of the NPS I&M Program, generally, and to address the logistical challenges of field work in SWAN, specifically.

2. Materials and methods

The initial steps of defining program goals and objectives, compiling and summarizing existing information, developing conceptual models, and prioritizing and selecting indicators (cf. Fancy et al., 2009), were precursors to the development of an overall sampling design for SWAN and are not addressed in this paper. In the following sections, we describe the process we used to develop a ground-based sampling design for monitoring vegetation in national parks in southwest Alaska.

2.1. Phase I: Defining a target population, sampling frame and sampled population

An important step in developing a sampling design for some geographical area is to define the spatial units containing the complete collection of plants, animals, or other natural resources of interest. The collection of all possible point locations or spatial units within this area of interest comprises the target population (Fig. 2). That is, one may define the target population as all point locations within an area where some spatial unit (e.g., plot) is associated with a selected point, or as a non-overlapping collection of spatial units that cover the area of interest. In our target population, we subdivide the area of interest into non-overlapping spatial units, or sampling units (e.g., plots), that do not occur within bodies of water, marine mudflats, airstrips, trails, cabins, campsites, and non-NPS lands within a given SWAN park. The natural resource of interest, such as a plant, is an attribute measured on the

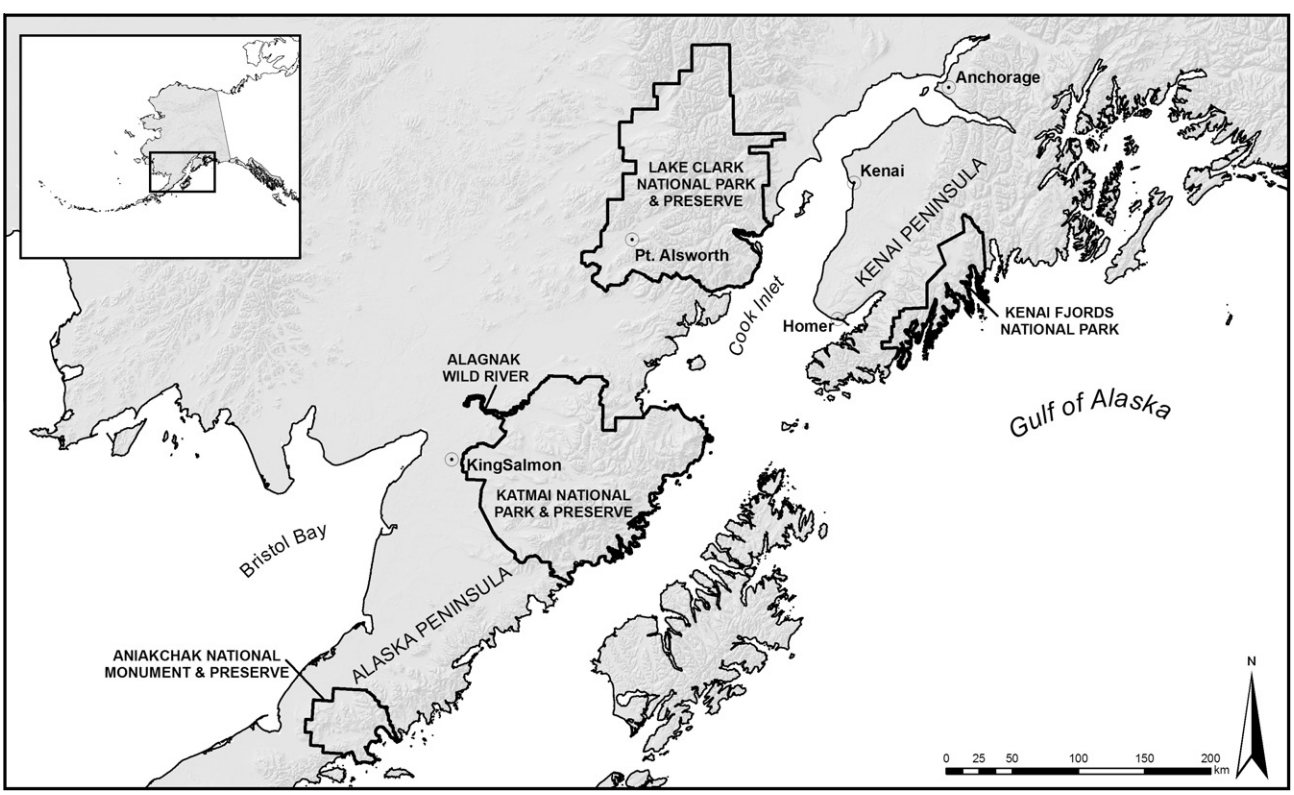

Fig. 1. The Southwest Alaska Network (SWAN) includes Alagnak Wild River, Aniakchak National Monument and Preserve, Katmai National Park and Preserve, Lake Clark National Park and Preserve and Kenai Fjords National Park. Collectively, these units sum to 3.8 million ha, or 11.6% of the land managed by the US National Park Service (figure from Bennett et al. (2006)).

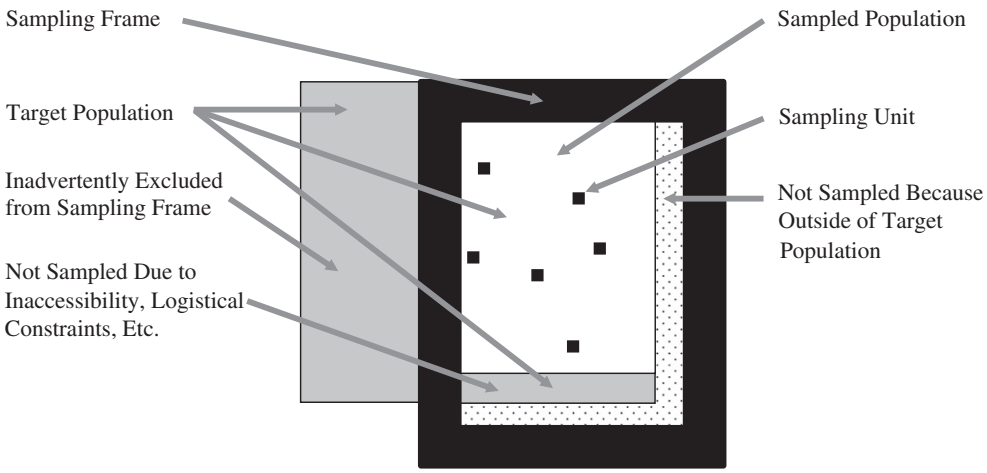

Fig. 2. Sampling terminology. Everything within the dark rectangle is part of the sampling frame, the combined gray and white areas represent the target population, the white area alone is the sampled population, and the stippled white area is outside of the target population. (Adapted from A.R. Olsen (unpublished presentation) and Lohr (1999)).

sampling units. This collection of sampling units forms the sampling frame. Ideally, a sampling frame will only contain the target population, but frame imperfections can lead to either exclusions of part of the target population (e.g., due to access limitations) or inclusions of portions of a non-target population (Fig. 2). The portion of the sampling frame that is available to be sampled is referred to as the sampled population. Thus, the geographic area containing the sampled population defines the spatial extent to which one can properly extrapolate results from a random sample of plots. Note that target and sampled populations are statistical entities that may or may not equate to populations in a strictly biological sense.

To address the challenges associated with access to sample areas, we used Path Distance analysis tools in ArcGIS (ESRI, Inc., Redlands, CA) to develop a GIS layer for identifying vegetated areas that could be safely reached by fixed-wing aircraft, boat, or foot in each park (Mortenson and Miller, 2008). These accessible areas, included in the resultant GIS layer, initially defined the sampled populations in the parks and were further refined during the next phase of our process (Section 2.2). Path Distance analysis provides a method for modeling the minimum cumulative travel cost from a starting point to surrounding locations within a specified area that adjusts for the actual surface distance by using horizontal (e.g., dense vegetation) and vertical (e.g., slope) factors that influence the total travel cost.

To create the access layer, we identified points of access through consultation with park staff. Starting points for foot travel included large lakes (>40 ha), airstrips, trails, and beaches where boat or float plane landings were possible. We buffered these source areas by 5000 m and converted the resultant layer to a raster data set (30 m cell size). We then removed the source points (e.g., cabins, airstrips, lakes were null areas buffered by 100 m) from the data set to avoid sampling in areas of human disturbance and/or open water. We ranked each remaining cell qualitatively by its relative cost to travel (0 = easy, 10 = impassable) in the backcountry. Cumulative cost rankings were derived from a combination of input data layers including (1) vegetation cover class (landcover classification; e.g., Golden and Spencer, 1998); (2) slope (Gesch et al., 2002); (3) walking distance from areas of access; (4) glaciers; (5) unwadeable rivers and streams; and (6) land ownership (see Mortenson and Miller, 2008 for ranks by data layer). A rank of 10 for any one of the input parameters rendered a cell impassible and it was removed from the access layer. For instance, slopes >50°, glaciers, non-NPS lands and unwadeable rivers were considered impassable. Cells with ranks <10 for individual parameters were included in the sample frame, although higher ranks (e.g., closed alder [*Alnus sp.*] stand, rank = 9) necessarily increased cumulative travel 'cost,' but were included if they occupied an area $\leqslant$90 m $\times$ 90 m (i.e., traversing through them over a short distance may be feasible). The final access layer with a cumulative travel cost assigned to each cell was attributed with slope, elevation, and landcover class and converted to a point coverage in a geodatabase that could be imported into programs SAS (SAS Institute, Inc., 2008) and S-Draw (http://www.west-inc.com/computer.php) for further analysis.

2.2. Phase II: Evaluating alternative sampling designs and refining the sampled population

The objectives of this phase were to use simulations to evaluate (1) the minimum level of change in a given attribute (e.g., landcover class) that could be detected when single random plot and random clusters of plots (i.e., 2 $\times$ 2 and 3 $\times$ 3 grids of plots) comprised the sample within each of three elevation bands (0–450 m; >450–900 m; >900 m); and (2) the sampling effort required to detect a minimum level of change when all landcover classes or only specific landcover classes comprised the sampled population. The cluster samples in (1) consisted of grids of plots of various spacing (30 m, 60 m, 120 m, and 240 m), where spacing was in multiples of the minimum resolution of the point coverage generated by the Path Distance analysis (Section 2.1). The largest spacing $\times$ grid size cluster combination represented the maximum area that could be feasibly sampled during a 1–2 day visit based on preliminary field surveys.

We ran the simulations based on points assigned within an access layer developed for Lake Clark National Park and Preserve (NPP). In these simulations, we used landcover classes assigned to each point in the access layer generated by the Path Distance analysis (Section 2.1) as surrogates for 30 $\times$ 30 m vegetation plots measured in the field. Thus, each single point or point in a cluster within our simulations represented a single vegetation plot used in field sampling. We applied a true change of 50% over a single time period because we expected 50% to be an ecologically meaningful level of change (e.g., in percent cover) for relatively common plant species at the plot level, and large enough in magnitude that it could be detected by observers over repeated sampling intervals. Points with slopes >10° were removed from the sampled population to minimize potential effects of erosion: i.e., digital elevation

models used to generate the slope values generally underestimated measured slopes in the field, so we used a conservative estimate of 10° to eliminate steeper areas (Section 4). We also excluded points with a landcover classification assigned as "unknown" because we required a spatial distribution of points with known landcover classifications for our simulations.

We used PROC SURVEYSELECT in SAS (SAS Institute, Inc., 2008) to randomly choose 50% of the points (plots) in two selected landcover classes (dwarf shrub tundra, prostrate shrub tundra) from three elevation bands (0–450 m; >450–900 m; >900 m) and to change their landcover classifications to one of the 23 remaining landcover classes between Times 1 and 2. The two landcover classes used in the simulations occurred across a range of frequencies (0.6–61%) within the three elevation bands in Lake Clark NPP. All remaining plots associated with other landcover classes were unchanged during the two time periods. The simulation therefore mimicked the situation wherein vegetation communities experience different rates of change because some are more sensitive to change than others.

We used the freeware program S-DRAW (T. McDonald, WEST, Inc.; http://www.west-inc.com/computer.php) to select 1000 generalized random-tessellation stratified (GRTS; Stevens and Olsen, 2004) samples of 50 single points (plots) from all landcover classes in accessible areas from each elevation band. Because of the fragmentary nature of the underlying point coverage and its apparent deleterious impact on spatial balance in the ordering of GRTS samples, we chose optimal pixel sizes in S-Draw by visual inspection of the spatial balance in the order of GRTS samples selected based on a range of pixel sizes for each elevation band. SAS (SAS Institute, Inc., 2006) was used to generate 2×2 and 3×3 clusters of plots of various spacing at each GRTS location. We used SAS PROC SURVEYMEANS (SAS Institute, Inc., 2008) to estimate the proportion of plots that changed landcover class between Times 1 and 2, and the associated 90% confidence interval, for each GRTS sample of $n = 2, 3, 4, \ldots, 50$ within each simulation run. The confidence intervals were based on variance estimators for simple random samples (single GRTS plots) and for cluster samples (grids attached to each GRTS plot), respectively; these variance estimators were very conservative approximations to the GRTS variance estimator, which was not available in SAS. We then computed the percentage of times that the lower 90% confidence limit of the estimated proportion was greater than and/or equal to some specified minimum detectable total change (MDTC), or observed level of change, for GRTS samples of $n = 2, 3, 4, \ldots, 50$ across all simulation runs for single plots and cluster samples of various spacing and configurations. This percentage represented the chance of detecting an MDTC of the specified magnitude based on a given GRTS sample size and true total change (i.e., 50%).

2.3. Phase III: Choosing a sample size and frequency sufficient to detect change

We used results from the Phase II simulations as the basis for a second round of simulations that evaluated the sample size and frequency required to detect a minimum change that would meet our monitoring objectives. Specifically, we conducted simulations to investigate the MDTC in vegetation cover of an individual species that could be detected with 95% confidence by various sample sizes ($n = 6$, 8, 12 and 24) at two different revisit frequencies (5 and 10 years) over a 31-year period for each combination of percent cover (average of 5%, 25% and 45%) and coefficient of variation (CV) of percent cover (5–50% by 5% increments). Preliminary simulations indicated that results were similar for percent cover values between 5% and 50%, so we used values of 5%, 25% and 45% as the true percent cover values and averaged their results after each run. We expected that a maximum of 32 plots could be measured in the field in any given year, and conjectured that a minimum of 6 plots per park $\times$ elevation band $\times$ vegetation class combination would be necessary for a reasonably precise trend estimator. We included two other factors of 24 (i.e., $n = 8$ and 12 plots) as additional sample size alternatives to evaluate via simulations.

Our second round of simulations was based on the discrete form of an exponential growth model for generating a trend in population size, $p_t = p_0\lambda^t$, in each simulation run, where p_t was the percent cover of an individual species at time t, p_0 was the initial percent cover value, λ was the finite population multiplier or $1 + R$ (R was the finite rate of population growth; Eberhardt, 1987). We used $\lambda = 0.978$, which corresponded to a -50% change over 31 years. We added process (temporal) variance (ε_t) to this model by randomly generating $\hat{\lambda}^t$ for each plot and year from a truncated normal distribution whose mean was the specified growth rate raised to the power of the relevant year t (i.e., λ^t) and whose standard deviation was based on the average standard error (SE) for three populations of Choco palm (*Astrocaryum mexicanum*; SE[λ] = 0.0162; Alvarez-Buylla and Slatkin, 1994), a perennial plant species for which published estimates of population growth rates were available. We incorporated sampling variation into each simulated time series by randomly generating a different initial estimated percent cover ($\hat{p}_0$; time 0) for each group of 24 plots from a truncated normal distribution whose mean was the specified value (5%, 25%, or 45%) and whose standard error (SE[p]) corresponded to the specified percent cover and CV (5–50% by 5% increments), where SE(p) = $p \times$ CV(p). The truncated normal distribution had a lower bound of 0 and an upper bound of 1. We used $\hat{p}_t = \hat{p}_0\hat{\lambda}^t$ to calculate the estimated proportions ($\hat{p}_t$) of each plot for years $t = 1$–31 in each simulated time series. Simulated values for each group of 24 plots were subsetted into 6, 8 and 12 plots and these four sample sizes were analyzed separately for each combination of temporal revisit frequency, true percent cover (averaged), and CV of percent cover. We assumed a complete census of individuals in each plot. The statistical program SAS (SAS Institute, Inc., 2006) was used to conduct all simulations.

Because we ultimately wished to estimate trends in percent cover separately for each park $\times$ vegetation class $\times$ elevation band combination, we simulated and analyzed time series from a single panel within the proposed rotating panel design (cf. McDonald, 2003; Table 1), noting that simulation results would similarly apply across all panels. The response variable in these models was the natural logarithm of estimated percent cover generated from simulations, and the predictor variable was year. We used SAS PROC GLIMMIX (SAS Institute, Inc., 2008) to fit a repeated measures, generalized linear model to each simulated time series for estimating change in percent cover and one-sided 95% confidence intervals. Note that 95% one-sided confidence intervals and 90% two-sided confidence intervals (used in the first round of simulations) share the same upper bound. We calculated percent annual change as (EXP[year coefficient] $- 1) \times 100\%$ and total percent change was calculated as $[(1 + \text{annual change})^{\text{years}} - 1] \times 100\%$ (Earnst et al., 2005). We applied the AICc model selection criterion (Burnham and Anderson, 2002) to choose the best-fitting covariance structure (i.e., variance components) for the repeated measures model based on a single randomly simulated time series of 32 plots per year for 31 years.

We then calculated the proportion of times out of 1000 that the upper one-sided 95% confidence interval of each trend estimate was either equal to and/or less than a range of total changes associated with a true change of -50% during a 31-year period for each combination of sample size, revisit frequency, and CV. We used pilot data from vegetation surveys in Lake Clark NPP (Miller, unpublished data), and rates of community change reported in the literature (Chapin et al., 2005; Boucher and Mead, 2006; Bowman

Table 1
Rotating panel design for monitoring vegetation in SWAN parks, in which a panel is a group of plots that is always sampled during the same year (McDonald, 2003). Each X represents 32 GRTS-selected plots, or four park × elevation band × vegetation class combinations. The panel design used in simulations evaluated both 5-year (below) and 10-year sample frequencies over a 31-year period rather than the 20-year period illustrated here. We expect to sample 1 panel per field season or 32 plots per year.

Year																				
Panel	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	X					X					X					X				
2		X					X					X					X			
3			X					X					X					X		
4				X					X					X					X	
5					X					X					X					X

et al., 2006; Cannone et al., 2007) to select a MDTC of −25%, assuming a true change of −50% (Sections 2.2 and 4). Simulation results were fitted to a logistic regression model (PROC LOGISTIC; SAS Institute, Inc., 2008) to generate predicted probabilities of detecting the user-specified MDTC across a range of CVs (0–50%) at 1% increments for graphically displaying simulation results.

In order to interpret simulation results as they related to our sampled population, we compared CVs derived from within-plot estimates of species cover in the field (Miller et al., 2009) to simulation results (i.e., estimated percent chance of detecting a MDTC at a specified level of true change) at specified CVs. Pilot data (e.g., species cover and frequency) from Lake Clark NPP were collected in fifteen, 30 m × 30 m vegetation plots over a two-year period and were used to estimate mean cover values and within- and among-plot variances for a range of species that varied in abundance across plots.

3. Results

3.1. Phase I: Sampling frame and initial sampled population

The access layer developed for Lake Clark NPP, based on Path Distance analysis, represented 11% of land area in the park (Fig. 3), and 7% of the vegetated area. Approximately 25% of the park was occupied by glaciers and/or water; the remaining area was either in private inholdings or other non-NPS lands (5%) or had no suitable landing areas in close proximity. Access layers developed for two other parks in SWAN represented 24% (Kenai Fjords NP) and 31% (Katmai NPP) of vegetated land area.

3.2. Phase II: Alternative sample designs and refined sampled population

The first round of simulations indicated that a 60-m spacing among plots within either a 2 × 2 or a 3 × 3 grid was the best at detecting change of the GRTS cluster designs we considered. However, regardless of the grid spacing, GRTS samples based on single plots typically required 20–50% fewer plots to detect a change in landcover classes than the equivalent number of plots in 2 × 2 GRTS cluster designs and 40–50% fewer plots than in 3 × 3 GRTS cluster designs.

When all landcover classes in each elevation band were included in the sampled population, the sample sizes required to detect a specified minimum change were estimated to be cost-prohibitive. Simulation results showed that a sample of ⩽40 single plots from a GRTS design or ⩽10 clusters from a GRTS design were

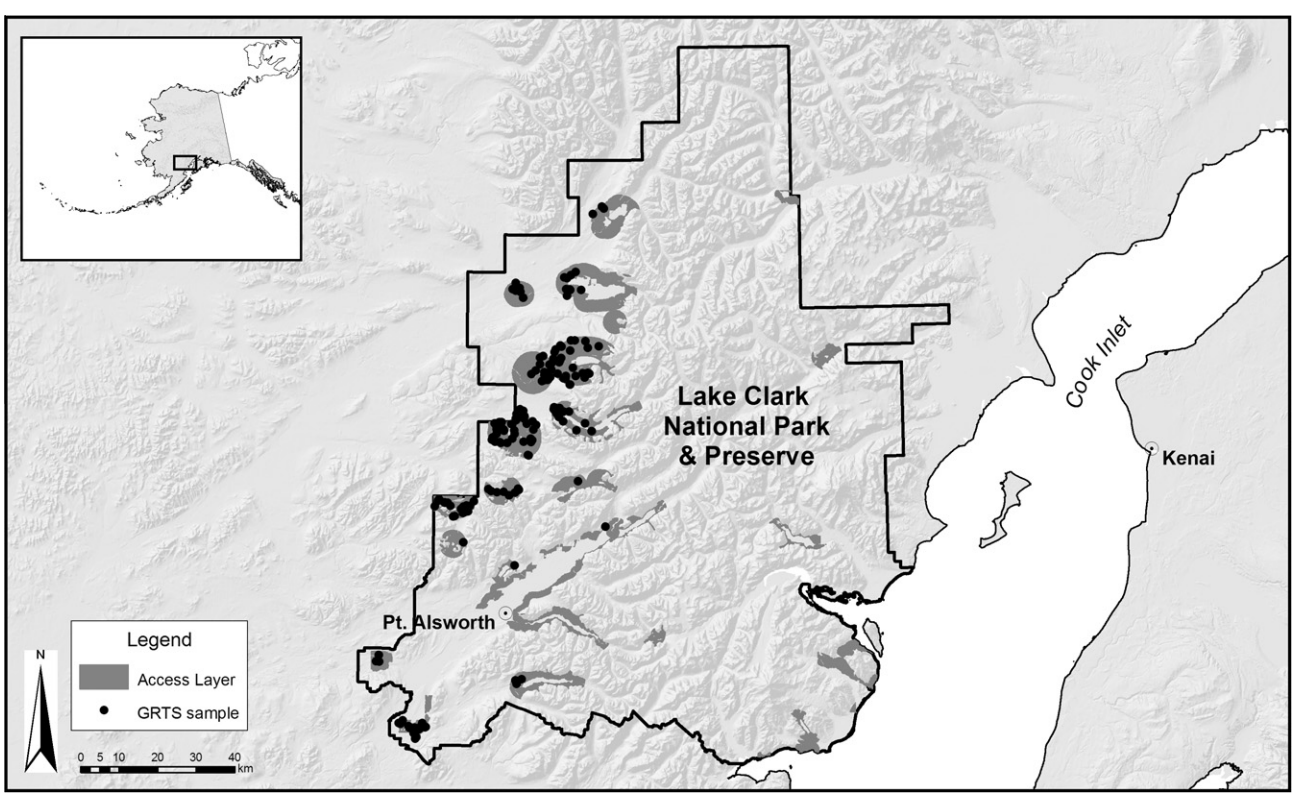

Fig. 3. Access layer for Lake Clark National Park and Preserve (gray) as generated via Path Distance analysis tools in ArcGIS. Within the access layer, a GRTS sample (black) was selected within each of three elevation bands. A subset of 200 GRTS points (plots) spanning the three elevation bands is shown here.

inadequate to detect at least a 25% change within the landcover classes of interest, given a true change of 50%, from a sampled population that included all landcover classes within each elevation band. Given a field season of 8–10 weeks with a minimum of two days of travel per sampling trip, we assumed that approximately 32 plots could be sampled over the course of a field season. The largest MDTC that could be detected under a realistic sampling regimen (30 single plots) was −15% for landcover classes that occupied at least 61% of an elevation band. In other words, we would need at least one field season to adequately sample a single, abundant (>60% cover) landcover class when sampling across all landcover classes, and this effort still would not allow us to meet our minimum change criterion of −25%.

The low ability to detect a minimum change associated with field-realistic sample sizes (e.g., ⩽40 plots) in a simulated sampled population that included all landcover types, indicated that we should restrict our sampled population in the field to specific landcover classes within each elevation band, as outlined by Woodward et al. (2009). Accordingly, we developed the second set of simulations (Section 2.3) applied to a sampled population consisting of a single landcover class × elevation band combination. Our preliminary selection of vegetation types for monitoring included low arctic tundra, boreal forest, alpine, and coastal forest. All of these vegetation types are relatively abundant on the landscape, are late-seral (Viereck et al., 1992), are of known ecological importance (e.g., as habitat), and are expected to be sensitive to a number of environmental drivers, including climate, pollutant loads, and disturbance. Collectively, they comprised nearly 60% of the total vegetated land area in the parks.

3.3. Phase III: Simulations to estimate sample size and frequency

The second round of simulations indicated that when the sampled population included a single landcover class × elevation band combination, the sample sizes required to detect a minimum change were much smaller than when sampling from all landcover classes within a given elevation band. Simulation results indicated we would have at least an 80% chance of detecting an MDTC of at least 25%, with 95% confidence, by sampling 8 plots per vegetation class × elevation band every 5 years when sample CVs of cover estimates in plots were 35% or less (Fig. 4). These results translate to 4 park × elevation band × vegetation class combinations that could potentially be sampled each year (total = 32 plots per year). Sample CVs for percent cover by species and growth form (e.g., shrub, forb, graminoid) were calculated from pilot data collected in Lake Clark NPP (Section 2.3) and compared with simulated CVs. Empirical CVs ranged from 7% to 79% for growth-form classes, and from 3% to >100% for selected species (Miller et al., 2009, 2010, available online). Widespread, relatively abundant species generally had sample CVs of ⩽35%, as did most growth-form classes, suggesting that the sample size estimated by simulations ($n = 8$) would provide us with sufficient ability to detect at least a −25% change in those individuals or classes in each park × elevation band × vegetation class over 31 years. Less-common or less-abundant species had sample CVs that were >35%, and thus appeared to be poor candidates for change detection.

4. Discussion

Logistical and budgetary constraints are inherent to any natural resource monitoring program, but they are especially significant in the large and remote national parks of Alaska. Our goal was to balance statistical inference with practical reality in a sampling design that could address several of the major challenges inherent to field studies in Alaska and elsewhere.

We used simulations and empirical data in a three-phased process to develop a sampling design that could be easily transferable to the field. First, we developed an access layer in a GIS to increase the probability of reaching sample plots in a reasonable period of time, and to initially demarcate the area of inference for the sampled population. Second, we conducted simulations to estimate the sample size required to detect an MDTC in a given landcover class when the sampled population included all landcover classes, and then refined our sampled population accordingly. Third, we used a separate round of simulations to estimate the sample size required to detect an MDTC in a given landcover class when the sampled population was constrained to specific landcover classes. We applied a decline to the population in this second round of simula-

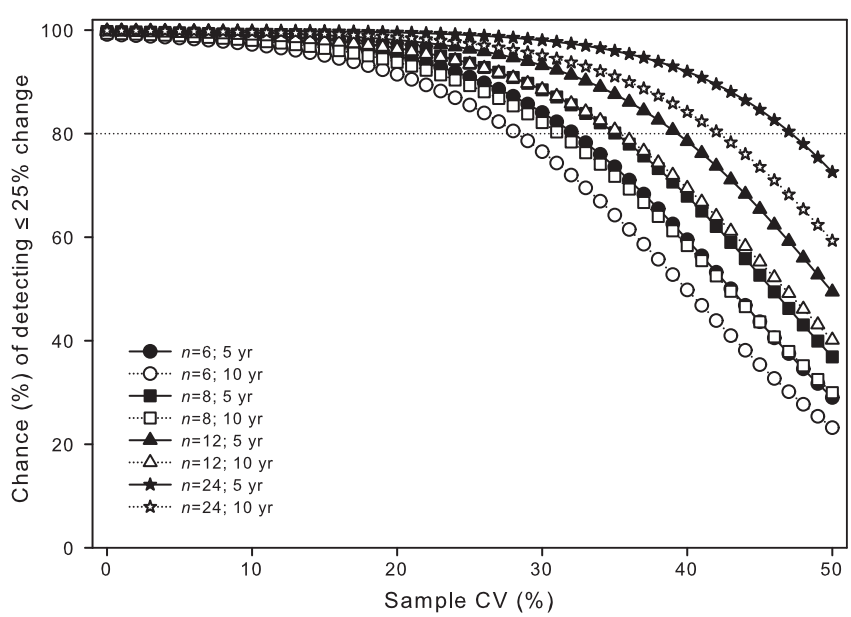

Fig. 4. Simulation results displaying the chance of detecting at least a −25% total change (MDTC) in vegetation cover, given a true change of −50%, with 95% confidence for a range of coefficients of variation (CV), sample sizes, and sample frequencies under a rotating panel design over 31 years. The horizontal dashed line indicates an 80% chance of detecting at least a −25% change.

tions, but we could have applied an increase – our results apply to either case. The process was iterative throughout: for instance, based on the simulation results, we modified the sampled population from all landcover types (Section 3.2) to individual landcover types (Section 3.3) to reduce sample variability and increase our ability to detect a minimum total change, thereby ensuring field-realistic sample sizes.

We assumed that climate change would be a major factor influencing vegetation change, which we used as a major criterion for defining the sampled populations as the vegetation types that were likely to be most widely affected by environmental drivers (e.g., climate) within each of three elevation bands across the SWAN parks. Our goal was to obtain estimates of change for each elevation band, which differs from a stratified design wherein sampling effort is allocated among strata to produce a single, overall estimate of trend. In any event, sampling across all vegetation types within each elevation band would add unnecessary noise to the change or trend estimators, which was borne out by results from our first set of simulations that indicated such an approach would not allow us to detect our minimum change criterion with field-realistic sample sizes. Sampling across all vegetation classes is effective for estimating status but is inefficient for estimating trend if classes vary in their frequency on the landscape, or if change occurs disproportionately across different classes, both of which are typically the case. Targeting specific vegetation classes requires that the area of inference (initial mapped boundaries of these classes) remains fixed throughout the monitoring program, regardless of future changes in the distribution of the classes.

Inaccuracies in the landcover classification and slope data led to imperfect definitions of our sampled populations (e.g., as represented by the gray area within the sampling frame displayed in Fig. 2). Landcover classification data for Lake Clark NPP had an overall misclassification error rate of approximately 17% (Golden and Spencer, 1998), whereas the 30-m digital elevation models used to generate slope data had an accuracy of no better than one-half of a contour interval, or 7.5 m (USGS, 2000). However, given that these errors are not systematic, the inferences drawn from our sampled populations should still be valid.

We selected a GRTS design over a two-dimensional, systematic random sample design with a random start because a GRTS design allows plots to be replaced from existing samples, such as those that are inaccessible or are not within the vegetation types of interest (e.g., as would occur with classification errors in the landcover map), while maintaining spatial balance. The GRTS design also has an appropriate variance estimator for reducing variance in the presence of spatial autocorrelation, which the traditional systematic sample design lacks. Simulation results indicated that a GRTS design based on single plots would be more effective at detecting change than a GRTS cluster design with the equivalent number of plots in a grid (e.g., single plots at four locations compared to a 2×2 cluster of plots at a single location). Although this comparison is skewed towards single plots because it does not include the cost savings of sampling adjacent plots in a grid, the gains in relatively lower sampling effort indicated in simulation results were large enough that we feel an adjusted comparison would still be in favor of single plots. This result is not surprising – it merely shows there was more variability in vegetation among sampling plots than within plots.

Our sampling objectives and simulations were based on an MDTC as a benchmark for detecting change rather than the conventional power-analysis approach based on the null hypothesis of no change. In other words, rather than estimating the chance of detecting change as the number of simulation runs in which a confidence interval or limit of a trend estimate did not include zero (i.e., a statistically significant result), we used the number of times these intervals or limits were at or below the MDTC (an ecologically meaningful result). This is an equivalence testing approach, but based on confidence intervals (Alderson, 2004), where the true total change is the effect size and the MDTC has been variously referred to as the expected effect size, minimum detectable biological difference, or minimum functional difference (Wellek, 2002; Brosi and Biber, 2009). In reality, our estimates of trend or change are subject to both temporal and sampling variation, so the change that can be detected is less than the true change that has occurred. A statistically significant result often is interpreted as the level of true change that occurred when, in fact, this is only true if the underlying model is true, and models are only approximations of reality. We prefer a more ecologically interpretable approach whereby a detectable change (e.g., −25%) is meaningful in and of itself for the most common species, and likely implies that a much larger true change (e.g., −50%) has occurred. Our approach also is more conservative than the traditional null hypothesis testing approach. For example, simulations evaluating sample size and frequency indicated that a minimum of 24 plots sampled every 5 years would be required to have at least an 80% chance of detecting an MDTC of −25% with 95% confidence, whereas a minimum of 6 plots every 5 years would be required based on an MDTC <0 (W. L. Thompson, unpublished results). We argue that it is better to oversample than to undersample. Although oversampling may waste some amount of a budget to achieve the monitoring objective, undersampling will waste an entire budget by failing to achieve this objective.

Our three-phased approach to developing a ground-based sampling design for monitoring vegetation uses the best available data in conjunction with simulations to produce a statistically defensible and logistically feasible protocol. We invested considerable time and effort into collecting data, conducting simulations, and thinking hard throughout this iterative process about which refinements to our design would enable us to best meet our monitoring objectives. This followed the initial steps of defining program goals and objectives, compiling and summarizing existing information, developing conceptual models, and prioritizing and selecting indicators (cf. Fancy et al., 2009). Such an investment at the initial stages of development of a sampling design will ultimately cost less than a lower initial effort resulting in a poorly designed protocol that does not meet monitoring objectives.

Acknowledgements

We thank A. Symstad, P. Geissler, E. Beever and two anonymous reviewers for their many helpful review comments that improved this manuscript. We also thank C. Roland for developing and sharing his sampling design that helped inform our work. Additional guidance came from A. Symstad, P. Spencer, T. Hollingsworth, W. Bowman, T. Jorgenson, and M. Carlson. Special thanks to C. Lindsay, C. Moore, and J. Walton for their valuable assistance in collecting field data. This work was supported by the National Park Service, SWAN I&M Program. Any use of trade names is for descriptive purposes only and does not imply endorsement by the US Government.

References

Alderson, P., 2004. Absence of evidence is not evidence of absence. British Medical Journal 328, 476–477.

Alvarez-Buylla, E.R., Slatkin, M., 1994. Finding confidence limits on population growth rates: three real examples revised. Ecology 75, 255–260.

Barber, V.A., Juday, G.P., Finney, B.P., 2000. Reduced growth of Alaskan white spruce in the twentieth century from temperature-induced drought stress. Nature 405, 668–673.

Bennett, A.J., Thompson, W.L., Mortenson, D.C., 2006. Vital signs monitoring plan, Southwest Alaska Network. National Park Service, Anchorage, Alaska. <http://science.nature.nps.gov/im/units/swan/index.cfm?theme=reports_pub>.

Berg, E.E., Henry, J.D., Fastie, C.L., De Volder, A.D., Matsuoka, S.M., 2006. Spruce beetle outbreaks on the Kenai Peninsula, Alaska, and Kluane National Park and Reserve, Yukon Territory: relationship to summer temperatures and regional differences in disturbance regimes. Forest Ecology and Management 227, 219–232.

Boucher, T.V., Mead, B.R., 2006. Vegetation change and forest regeneration on the Kenai Peninsula, Alaska following a spruce beetle outbreak, 1987–2000. Forest Ecology and Management 227, 233–246.

Bowman, W.D., Gartner, J.R., Holland, K., Wiedermann, M., 2006. Nitrogen critical loads for alpine vegetation and terrestrial ecosystem response: are we there yet? Ecological Applications 16, 1183–1193.

Brosi, B.J., Biber, E.G., 2009. Statistical inference, Type II error, and decision making under the US Endangered Species Act. Frontiers of Ecology and the Environment 7 (9), 487–494.

Burnham, K.P., Anderson, D.R., 2002. Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach. Springer-Verlag, New York.

Cannone, N., Sgorbati, S., Guglielmin, M., 2007. Unexpected impacts of climate change on alpine vegetation. Frontiers of Ecology and the Environment 5 (7), 360–364.

Chapin III, F.S., Sturm, M., Serreze, M.C., McFadden, J.P., Key, J.R., Lloyd, A.H., McGuire, A.D., Rupp, T.S., Lynch, A.H., Schimel, J.P., Beringer, J., Chapman, W.L., Epstein, H.E., Euskirchen, E.S., Hinzman, L.D., Jia, G., Ping, C.-L., Tape, K.D., Thompson, C.D.C., Walker, D.A., Welker, J.M., 2005. Role of land-surface changes in Arctic summer warming. Science 310, 657–660.

Earnst, S.L., Stehn, R.A., Platte, R.M., Larned, W.W., Mallek, E.J., 2005. Population size and trend of yellow-billed loons in northern Alaska. Condor 107, 289–304.

Eberhardt, L.L., 1987. Population projections from simple models. Journal of Applied Ecology 24, 103–118.

Fancy, S.G., Gross, J.E., Carter, S.L., 2009. Monitoring the condition of natural resources in US national parks. Environmental Monitoring and Assessment 151, 161–174.

Field, S.A., O'Connor, P.J., Tyre, A.J., Possingham, H.P., 2007. Making monitoring meaningful. Austral Ecology 32, 485–491.

Gesch, D., Oimoen, M., Greenlee, S., Nelson, C., Steuck, M., Tyler, D., 2002. The National Elevation Dataset. Photogrammetric Engineering and Remote Sensing 68, 5–11.

Golden, M., Spencer, P., 1998. Lake Clark National Park and Preserve land cover mapping project user's guide. Natural Resource Technical Report NPS/LACL/NRTR—1998/001. National Park Service, Fort Collins, Colorado. <http://www.nature.nps.gov/publications/NRPM/nrtr.cfm>.

Hudson, J.M.G., Henry, G.H.R., 2009. Increased plant biomass in a High Arctic heath community from 1981 to 2008. Ecology 90, 2657–2663.

Legg, C.J., Nagy, L., 2006. Why most conservation monitoring is, but need not be, a waste of time. Journal of Environmental Management 78, 194–199.

Lesica, P., McCune, B., 2004. Decline of Arctic-alpine plants at the southern margin of their range following a decade of climatic warming. Journal of Vegetation Science 15, 679–690.

Lindenmayer, D.B., Likens, G.E., 2009. Adaptive monitoring: a new paradigm for long-term research and monitoring. Trends in Ecology and Evolution 24, 482–486.

Lohr, S.L., 1999. Sampling: Design and Analysis. Duxbury Press, Pacific Grove, California.

McDonald, T.L., 2003. Review of environmental monitoring methods: survey designs. Environmental Monitoring and Assessment 85, 277–292.

Miller, A.E., Thompson, W.L., Moore, C., 2009. Vegetation composition and structure: baseline monitoring in the Southwest Alaska Network, 2008 annual summary report. Natural Resource Technical Report NPS/SWAN/NRTR—2009/213. National Park Service, Fort Collins, Colorado. <http://science.nature.nps.gov/im/units/swan/Libraries/Reports/SWAN_Veg_2008_Annual_Report_FINAL_090707.pdf>

Miller, A.E., Walton, J.K., Moore, C., Thompson, W.L., 2010. Vegetation composition and structure: baseline monitoring in the Southwest Alaska Network, 2009 annual summary report. Natural Resource Technical Report NPS/SWAN/NRTR—2010/342. National Park Service, Fort Collins, Colorado. <http://science.nature.nps.gov/im/units/swan/Libraries/Reports/MillerA_2010_SWAN_VegCompStrucBaselineAnnualReport2009NRTR_20100610.pdf>

Mortenson, D.C., Miller, A.E., 2008. Path Distance analysis to determine accessibility for long-term vegetation monitoring program, Lake Clark National Park and Preserve. Natural Resource Technical Report NPS/SWAN/NRTR—2008/126. National Park Service, Fort Collins, Colorado. <http://science1.nature.nps.gov/naturebib/biodiversity/2008-9-25/MortensonD_2008_PathCostAnalysis_652486.pdf>

SAS Institute, Inc., 2006. Base SAS® 9.1.3 Procedures Guide, Vols. 1–4, second ed. SAS Institute, Inc., Cary, North Carolina.

SAS Institute, Inc., 2008. SAS/STAT® 9.2 User's Guide. SAS Institute, Inc., Cary, North Carolina.

Stevens Jr., D.L., Olsen, A.R., 2004. Spatially balanced sampling of natural resources. Journal of the American Statistical Association 99, 262–278.

Sturm, M., Racine, C., Tape, K., 2001. Climate change: increasing shrub abundance in the Arctic. Nature 411, 546–547.

USGS, 2000. US GeoData digital elevation models. US Geological Survey, EROS Data Center, Fact Sheet 040-00, Sioux Falls, South Dakota. <http://egsc.usgs.gov/isb/pubs/factsheets/fs04000.pdf>

Viereck, L.A., Dyrness, C.T., Batten, A.R., Wenzlick. K.J., 1992. The Alaska vegetation classification. US Forest Service, Pacific Northwest Region, General Technical Report PNW-GTR-286, Portland, Oregon.

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, O., Turner, P.L., Tweedie, C.E., Webber, P.J., Wookey, P.A., 2006. Plant community responses to experimental warming across the tundra biome. Proceedings of the National Academy of Science 103, 1342–1346.

Wellek, S., 2002. Testing Statistical Hypotheses of Equivalence. CRC Press, Boca Raton, Florida.

Woodward, A.E., Hutten, K.M., Boetsch, J.R., Acker, S.A., Rochefort, R.M., Bivin, M.M., Kurth, L.L., 2009. Forest vegetation monitoring protocol for national parks in the North Coast and Cascades Network. US Geological Survey Techniques and Methods 2-A8, Reston, Virginia. <http://pubs.usgs.gov/tm/tm2a8/>.