Biological Conservation 144 (2011) 1278–1284

Contents lists available at ScienceDirect

Biological Conservation

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

Planning for success: Identifying effective and efficient survey designs for monitoring

Joel H. Reynolds [a,*], William L. Thompson [b], Brook Russell [a,c]

[a] US Fish & Wildlife Service, Division of Realty and Natural Resources, 1011 E. Tudor Road, MS 221, Anchorage, AK 99503, United States
[b] US National Park Service, Southwest Alaska Network, 240 West 5th Avenue, Anchorage, AK 99501, United States
[c] Casper College, Dept. of Statistics, 125 College Drive, Casper, WY 82601, United States

ARTICLE INFO

Article history:
Received 11 April 2010
Received in revised form 16 November 2010
Accepted 1 December 2010
Available online 3 January 2011

Keywords:
Cost analysis
Monte carlo simulation
Population trend
Sample size
Statistical power
Study planning

ABSTRACT

Selecting a survey design to detect change through time in an ecological resource requires balancing the speed with which a given level of change can be detected against the cost of monitoring. Planning studies allow one to assess these tradeoffs and identify the optimal design choices for a specific scenario of change. However, such studies seldom are conducted. Even worse, they seem least likely to be undertaken when they offer the most insight – when survey methods and monitoring designs are complex and not well captured by simple statistical models. This may be due to limited technical capacity within management agencies. Without such planning, managers risk a potentially severe waste of monitoring resources on ineffective and inefficient monitoring, and institutions will remain ignorant of the true costs of information and the potential efficiency gains afforded by a moderate increase in technical capacity. We discuss the importance of planning studies, outline their main components, and illustrate the process through an investigation of competing designs for monitoring for declining brown bear (*Ursus arctos*) densities in southwestern Alaska. The results provide guidance on how long monitoring must be sustained before any change is likely to be detected (under a scenario of rather strong true decline), the optimal designs for detecting a change, and a tradeoff where accepting a delay of 2 years in detecting the change could reduce the monitoring cost by almost 50%. This report emphasizes the importance of planning studies for guiding monitoring decisions.

Published by Elsevier Ltd.

1. Introduction

An ecological monitoring program created to inform management decision-making is only *effective* if it produces information of sufficient accuracy and precision as to resolve the underlying scientific questions (Field et al., 2007; Lindenmayer and Likens, 2009) and thus influence the decision-making (Olsen et al., 1999; Lyons et al., 2008). When an underlying question entails detecting a specified level of population change, perhaps in response to a specific management action or as part of a wider investigation into ecosystem change, then the faster the monitoring program can detect the given level of change when it indeed happens the more effective the program will be. However, limited monitoring resources (staff, money, equipment) constrain effectiveness by also requiring *efficiency;* the goal is a monitoring program that provides the quickest detection for the least implementation cost (Sims et al., 2008).

A large determinant of both a monitoring program's effectiveness and its efficiency is the survey design - the distribution of data collection effort in space and time. Planning studies allow one to identify combinations of design choices, especially the per-survey sampling effort level and survey frequency, that provide optimal tradeoffs in effectiveness and efficiency (Rhodes et al., 2006, provide an alternative formulation).

The general importance of such planning is widely acknowledged in the literature's discussion of statistical power analysis (Gerrodette, 1987; Thomas and Krebs, 1997; Sims et al., 2006), yet it remains the exception rather than the rule in natural resource monitoring (Legg and Nagy, 2006; Field et al., 2007; Lindenmayer and Likens, 2009). Unfortunately, failure to conduct such planning can entail severe management costs. A program's ineffectiveness may not be recognized and remain unresolved for a long time given the large spatial and temporal variability of most resources of interest in natural systems. Meanwhile, program resources (staff time, money, equipment) are wasted that could have been directed elsewhere (e.g., opportunity costs); the poor-quality information delays recognition of and response to important system changes, reducing management capacity and efficiency because the changes have progressed further before detection. This may lead to misguided management decisions and irreversible shifts in the ecosystem (Fairweather, 1991; Taylor and Gerrodette,

* Corresponding author. Tel.: +1 907 786 3914; fax: +1 907 786 3905.
E-mail addresses: joel_reynolds@fws.gov (J.H. Reynolds), bill_thompson@nps.gov (W.L. Thompson), brooktrussell@yahoo.com (B. Russell).

0006-3207/$ - see front matter Published by Elsevier Ltd.
doi:10.1016/j.biocon.2010.12.002

1993; Gibbs et al., 1999; Reid, 2001; Field et al., 2005; Legg and Nagy, 2006; Taylor et al., 2007).

Long-term costs of poor planning include failing to adequately document historic ecosystem processes and population levels to aid future planning, and wasting resources by failing to improve the monitoring process itself (Nichols and Williams, 2006; Lyons et al., 2008). From an organizational perspective, the most important shortcomings may be the failure to learn both how to develop well-defined management objectives and the technical aspects of developing well-designed monitoring programs (Field et al., 2007). This raises the subsequent risk of losing long-term institutional support. All of this is in addition to the cumulative financial costs, which can be substantial, of repeatedly spending 'trivial' amounts to monitor X.

Although perhaps the less frequent problem, a monitoring program is inefficient if it employs excessive levels of survey effort and frequency relative to the information quality required for the decision-making. Such programs similarly entail cumulative costs of implementation, opportunity costs, and perhaps greater risk of losing institutional support due to the potential savings from stopping the program.

The importance of planning increases with survey costs and potential for political contention over survey results. Both features characterize current and proposed landscape-scale monitoring efforts focused on understanding and responding to global change in northern latitudes (for example, see Beever and Woodward, in this issue). The complexity, and value, of the planning process increases with the observation process's complexity, such as known major sources of bias (e.g., imperfect detection), multiple sources of variation (Renner et al., 2010), and/or logistical constraints (see Thompson et al., in this issue). All of these features characterize the monitoring of large mammals in remote regions of Alaska, especially brown bears (*Ursus arctos*) (e.g., Walsh et al., 2010). Brown bear surveys are expensive, logistically challenging, have a high potential for political contention over the results, and can require complex survey methods.

Finding an optimal, or near optimal, combination of per-survey effort and survey frequency, i.e., one that achieves the required information quality standards for the lowest cost, can be done directly if the number of feasible combinations is small. Each combination is evaluated for its cost and ability ('statistical power') to detect the pre-defined level of change or trend at the desired level of statistical significance. We illustrate the planning process by investigating the effort levels and costs required to detect declines in brown bear density on Togiak National Wildlife Refuge (NWR) in Alaska, USA.

1.1. Is it feasible to monitor brown bears on Togiak National Wildlife Refuge?

Brown bears serve key ecological roles in Togiak NWR, a relatively undeveloped region of southwest Alaska (Walsh et al., 2010). They are top predators that influence population dynamics of other species and act as conduits of nutrient transfer from spawning salmon (*Oncorhynchus* spp.) to the terrestrial system (Gende et al., 2002). Brown bear management is important in this region because of the species role in the ecosystem and concerns over its impact on subsistence prey species, such as caribou and moose (Walsh et al., 2010). Precise estimates of brown bear density are needed for immediate management decisions regarding harvest regulations, whereas precise estimates of changes in density are needed for properly interpreting and reacting to changes in prey species abundance (National Research Council, 1997). However, brown bears are difficult and expensive to survey in southwest Alaska (Becker and Quang, 2009). They often occur at low to moderate densities, can have low probability of detection from aircraft except in open habitats, are inactive throughout the winter, and often inhabit areas that are largely remote, mountainous, and/or difficult to access (Walsh et al., 2010).

Mark-recapture methods using radio collars have been employed to estimate brown bear abundance in the Togiak region (Kovach et al., 2006) but these methods were not sustainable as a regular monitoring program due to their required effort and costs. DNA mark-recapture using hair snares (Boulanger et al., 2002) would eliminate the need for direct handling of bears but still suffers from the logistical constraints and expense associated with ground-based efforts to establish and maintain hair snares (ibid; Boulanger et al., 2006) in a region dominated by mountains and under the operating constraints of legally protected wilderness. A line transect method using distance sampling from small aircraft and double-observer models was developed to survey brown bears (Becker and Quang, 2009) and successfully used in other parts of southwest Alaska (ibid). The method does not require ground access or direct handling of bears. However, experience suggests it requires approximately 150 or more detections of bear groups to attain relatively precise abundance estimators (Earl Becker, Alaska Dept. of Fish and Game, pers. comm.).

This distance sampling method was employed in the Togiak region in 2004/2005, the second year being required to achieve the minimum of 150 bear group detections (see Table 2 in Walsh et al., 2010). In total, 969 25-km long transects were surveyed and 199 bear groups detected, leading to an estimate of 40.3 bears per 1000 km^2 (95% confidence interval 31.4–54.5) (Walsh et al., 2010). The data collection required 20 survey days each for five survey aircraft with two-person crews (front observer/pilot, back observer/record keeper); a breakdown of implementation costs by task is in the online supplement of Walsh et al. (2010).

Given the successful implementation of the survey in 2004/2005, the manager of Togiak NWR raised two questions motivating the planning study reported here: What survey effort levels and frequency would most efficiently achieve a probability of 0.80 of detecting a population rate of decline meeting the IUCN 'vulnerable' criterion (IUCN Standards and Petitions Subcommitee, 2010)? What would that program cost?

We summarize the planning steps used to address the Togiak NWR manager's questions. We emphasize the general planning components as a broad guide for resource managers to the steps involved, minimizing the technical details of this particular application. Our objective is to illustrate to managers (i) the value of such planning and (ii) the relatively negligible cost of conducting such planning relative to the total cumulative cost of implementing a monitoring program.

2. Methods

The manager's questions were addressed by calculating, for each survey design considered, the smallest rate of decline detectable (the *minimum detectable difference*) with the manager's specified statistical error rates; the minimum detectable difference was calculated after each survey through the first 41 years of the monitoring program. A survey design became effective the year its minimum detectable difference was smaller than the rate of population change associated with classifying the population as 'vulnerable' under IUCN criteria (IUCN Standards and Petitions Subcommittee, 2010). This type of planning study requires defining a number of features before implementation.

2.1. Define the survey design options

All combinations of per-survey sample size (*n* = 500, 750, 1000, 1250 or 1500 transects) and survey frequency (every 1, 4, 5, or

10 years) were investigated, with one exception. Results from annual monitoring were only considered for $n = 250$ or 500 transects because it is logistically infeasible to sample more than 500 transects in a given season. This constraint reflected limitations in (i) the time available between den emergence and vegetation green-up (Walsh et al., 2010), (ii) the availability of experienced pilots and observers to conduct such studies, (iii) annual agency budgets for biological surveys, and (iv) safety and logistical concerns of simultaneously employing more than five survey aircraft in the study region.

2.2. Define the survey cost function

Cumulative costs of implementing each survey were estimated based on the costs of the 2004/2005 Togiak NWR survey: fixed costs for initiating monitoring = US$9000.00, fixed costs per survey = US$21,460.00, and per transect costs = US$168.51 (Patrick Walsh, Togiak NWR, pers. comm.). Calculations included a 3% per annum inflation rate across the duration of monitoring. The costs accounted for each survey's data processing, QA/QC, and analysis but did not include long-term data management or dissemination of survey results.

2.3. Define the response metric

Given the limited information on brown bear populations in this region, the initially proposed monitoring objective focused on estimating instantaneous rates of population change, r, assuming an exponential growth model with multiplicative errors (*Process model*, Table 1). This could be estimated using just the first and most recent surveys, $\hat{r} = \ln(N_t/N_0)/t$, or using the slope estimate from fitting a regression model of abundance estimates from all the surveys against time (Skalski et al., 2005). The latter approach utilized more information so was assumed to provide more precise rate estimates. Accordingly, rate of decline was estimated by fitting the linear model of $\ln(\widehat{N_t})$ against survey year ('t') using weighted least squares to account for any changes in the estimated variance of the $\ln(\widehat{N_t})$ values as true density declined. Weights were inversely proportional to the var(ln $(\widehat{N_t})) \approx \frac{\widehat{var(N_t)}}{(\widehat{N_t})^2}$ by Taylor series expansion.

2.4. Define the reference rate of change and statistical error rates

The reference rate of decline was set at the value required for classifying the population as 'vulnerable' under IUCN criteria (IUCN Standards and Petitions Subcommittee, 2010): an estimated 30% reduction in population size over the last 10 years or three generations, whichever is longer, where the reduction or its causes may not have ceased or may not be understood or may not be reversible. Generation length, defined to be the average age of the parents of the current cohort, was set at 16.5 years (Steve Kovach, pers. comm.; David Garshelis, pers. comm.). Under the exponential growth model, this $r_0 = \ln(N_t/N_0)$/(three generations) = ln(0.70)/49.5 = −0.0072. This rate of population change, sustained over 49.5 years, would reduce the population by 30%. Detecting a rate of population change of this magnitude or more extreme, though over a shorter time frame, was taken as a trigger by the manager. The comparison of the estimated rate of population change and this reference value used a Type I error rate of 0.05 and a Type II error rate of 0.20 (i.e., statistical power of 0.80).

Table 1
Process and observation models used to generate time series of density estimates and their standard errors, and the estimation model used to estimate the rate of population change from the each simulated time series. Process model parameter settings are given in Section 2.7. Source of the observation model is explained in Section 2.5.

Model	Components
Process	$N_t = N_{t-1}\exp(r + \varepsilon_t)$
	ε_t i.i.d. $\sim$ Normal(0, σ^2_{annual})
Observation	$\widehat{D_t} \sim \text{Normal}(D_t, \sigma^2_{\text{Density}}(D_t, n))$
	$D_t = N_t/\text{area}$
	$\sigma^2_{\text{Density}}(D_t, n) = (129.29 + 1.81D_t)^2/n$
	$\log(\widehat{SE}(\widehat{D_t})) \sim \text{Normal}(\quad)$
	$\mu(D_t, n) = 1.21 + 0.137\sqrt{D_t} - 0.000394n$
	$\sigma^2_{SE}(D_t, n) = 246.17 - 44.36\log(D_t)/n$
Estimation	$\log(\widehat{N}_t) = \beta_0 + rt + \delta_t$
	δ_t i.i.d. $\sim$ Normal(0, $\sigma^2_{\text{residual}}$)

Table 2
Optimal choices for survey effort levels and monitoring frequency change from among those investigated (Section 2.1), to most quickly and least expensively detect a rate of population change that, if sustained long enough, would warrant classifying the population as 'vulnerable' under the IUCN criterion, with a probability of 0.80 when the true rate of population change is $r = -0.03$, for each investigated initial density. 'True decline' is 1 − (mean of simulated true densities at that year)/(initial density).

Initial density (bears/1000 km^2)	First year	True decline (%)	Cumulative cost to first year (US$)	Monitoring frequency	Survey sample size
40.3	29	58	3,661,054	4	1500
	31	61	1,918,775	10	1500
80	16	38	1,256,863	5	1250
	21	47	670,439	10	750
160	11	28	717,953	5	1000
	21	47	490,360	10	500

2.5. Calculate the minimum detectable rate of decline for each design

Planning requires calculating how the variance of the estimate changes across different survey designs. When the quantity of interest is a rate or trend estimate, this requires calculating both how the per-survey estimate variance changes with sample size and how the trend estimate's variance changes with survey frequency. This can be calculated directly when the estimation methods allow derivation of the necessary analytical formulas (e.g., Larsen et al., 2001), otherwise calculations must use simulation.

As with many wildlife survey methods, the distance sampling method for brown bear density estimation combined a classical statistical survey design method, e.g., Horwitz–Thompson estimation, with fitting a probability model of the observation process, e.g., a detection function (Becker and Quang, 2009). The detection function was moderately complex as it included a general linear model to allow parameters of the detection function to systematically vary with covariates and a double-observer mark-recapture model to estimate maximum probability of detection (ibid). This complexity precluded deriving an analytical formula for how the variance of the density estimate would change as a result of changes in the sample size and underlying true density. Thus planning required simulation.

Simulation mimicked both the dominant sources of variation in the underlying resource of interest and the assumed change in that resource (via a *process* or *state model*) as well as the dominant sources of variation in surveying the simulated resource (via an *observation model*) (Borchers et al., 2002). The procedure had four steps (detailed below): (i) use the process model to generate a time sequence of declining true densities in the study region, (ii) use the observation model to generate a time sequence of density estimates and standard errors based on the true densities and the sample size, (iii) extract the estimates for the years actually surveyed under the specific survey frequency, then (iv) use those density estimates and their standard errors to estimate the response metric described above (Section 2.3).

2.6. Process model

Time series of declining brown bear abundance were generated over a 41-year period from a discrete-time exponential growth model with annual process variation imposed on the population growth rate (Table 1). Each abundance time series was converted to a density time series by dividing by the area of the simulated study region.

2.7. Observation model

The observation model was applied to each density time series to generate a time series of estimated densities and their uncertainties. The observation model included both the imperfect detection process during sampling and the distance sampling estimation process required by the unbiased density estimators. It was developed in two stages. Initially, computer code was written to simulate explicitly each step of surveying n transects in a simulated study region with a true brown bear density D and analyze the results (Appendix A). This model was too slow to directly employ in the trend simulations, so a small study was conducted to develop analytical models (Appendix B) that approximated the sampling distributions of both the density estimate, $\hat{D}$, and its standard error, $\widehat{SE}(\hat{D})$, as predicted by the detailed model, for a given true density and sample size (Table 1). The secondary models were then used to simulate brown bear density estimates and their standard errors in the planning study.

For each simulated time sequence of density estimates and standard errors generated, the relevant subset of estimates was extracted for each survey frequency and used to estimate the rate of population change, $\hat{r}$, and its 90% confidence interval at years 6, 11, 16, 21, 26, 31, 36 and 41. This procedure was repeated 1000 times for each sample size for each initial density. The 0.80 quantile of each set of 1000 upper confidence interval limits was that survey design's minimum detectable rate of decline. A decline <0 corresponded to having at least 80% power for rejecting a null hypothesis of no trend vs. the alternative hypothesis of declining trend.

2.8. Identification of optimal designs

For each initial density, the optimal survey designs were identified by plotting the first year at which a design's minimum detectable rate of decline was more extreme than, i.e. below, the IUCN threshold (Section 2.4) against the cumulative cost of implementing the design through that year. The optimal designs are those that had the quickest detection for a given total cost such that any design that detected the decline more quickly had a larger total cost. These optimal designs fall along the edge of the scatter near the horizontal and vertical axes, i.e., the tradeoff frontier.

2.9. Planning study settings

Population trends were simulated for initial densities (D_0) of 40.3, 80 and 160 brown bears per 1000 km^2, a range encompassing estimates from both the Togiak NWR study (Walsh et al., 2010) and a recent study from other southwest Alaska federal management units, Lake Clark National Park and Preserve and Katmai National Park and Preserve (Olson and Putera, 2007). The study area was a square 145.53 km on a side, giving the same area as the Togiak NWR study (21,178 km^2). The true instantaneous rate of population change was set at $r = -0.03$. The annual process variation around this trend was set at $\sigma_{\text{annual}} = 0.008$ based on a mark-recapture study from the Togiak NWR area (Kovach et al., 2006).

3. Results

Under the assumed process model, at an initial brown bear density as low as that estimated to exist at Togiak NWR (40.3 bears per

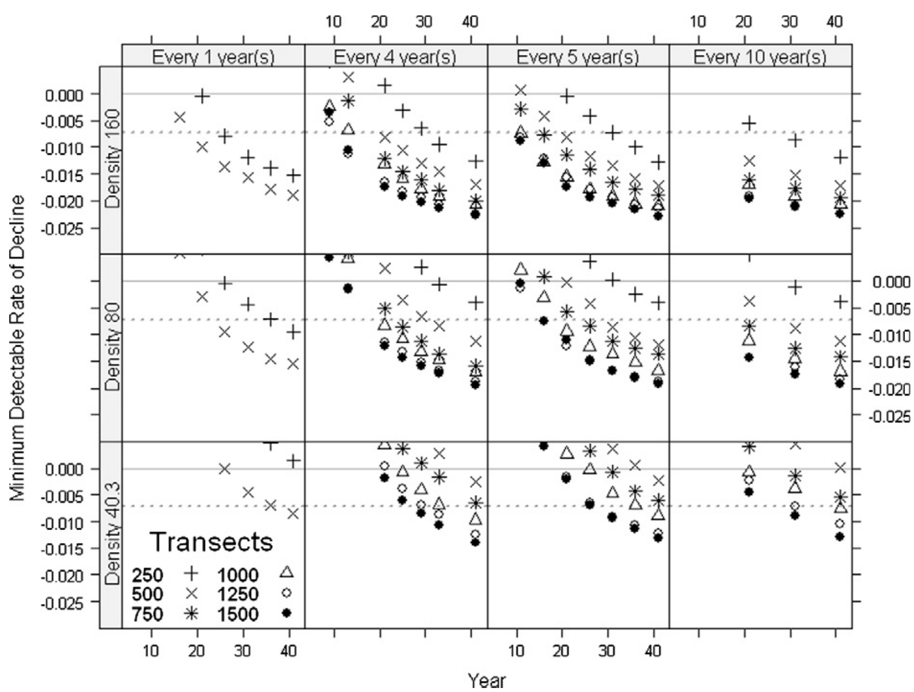

Fig. 1. Decrease in minimum rate of population change (vertical axis) detectable with probability of 0.80 as monitoring progresses through time (horizontal axis), for each initial density (row) and design combination (columns and symbols, see key in lower left panel). See Section 2.5 for details. A point below the solid horizontal line has a power of at least 0.80 to detect a declining population; a point below the dashed horizontal line has at least that power to detect a rate of change that would meet the IUCN vulnerable population criteria if sustained for three generations, e.g., 49.5 years.

1000 km^2), 21 years passed before any design combination detected a population decline (n = 1500 transects, survey every 4 years; minimum detectable rate of decline <0, Fig. 1) with the required probability of 0.80, even with the highest level of sampling effort; at this point the population was expected to have declined by at least 47%. At the earliest, managers should expect to monitor 29 years to attain their information objective of detecting a rate of population decline as bad as that required to meet the IUCN vulnerable criterion (though the rate of decline had not yet been sustained for three generations), surveying 1500 transects every 4 years (Fig. 1) at a total cumulative cost of over US$3.5 million (Table 2). The population would have actually declined, on average, 58% (Table 2). This level of monitoring effort would require surveying in three of every 5 years, assuming logistic constraints limit surveying to 500 transects a year. Or the management information objective could be met 2 years later at just over 1/2 the cost by surveying 1500 transects every 10 years (Table 2).

The optimal monitoring designs differed for different initial densities, with the optimal sample size decreasing as initial density increased (Fig. 1, Table 2). The management information objective was achieved much more quickly and cheaply at higher initial densities (Table 2).

4. Discussion

Planning studies are useful both for what they reveal and what they encourage. The current study revealed that Togiak NWR can achieve its initially proposed management information objectives by monitoring brown bears using the distance sampling survey method if it is willing to commit to spend approximately US$2 M over the next 30 years and survey 500 transects for three consecutive years per decade (Table 2). Although these results likely somewhat overestimate the required time and cost of achieving these information objectives (discussed below), managers can use them as a guide for assessing whether this is a feasible expenditure of monitoring resources and, if it is, to identify a design that is both effective and efficient.

More importantly, the results encourage a greater dialogue between managers and field biologists (and biometricians) regarding appropriate and feasible management objectives, potential management actions and information needs (including appropriate error rates). Such a dialogue may lead to selection of a different response metric or perhaps just different error rates. Setting Types I and II error rates is a policy choice and should reflect both the knowledge base of the relevant body of science and the risks associated with the alternative actions stemming from the potential decisions (Shrader-Frechette and McCoy, 1992; Field et al., 2004; Taylor et al., 2007). Given the limited information on brown bears in this region, it may be adequate to use a Type I error rate of 0.10 or 0.20, thus equalizing the risks of false discovery and missed signals (Types I and II errors, respectively). The planning study could then be recalculated to determine the optimal designs, expected monitoring duration and total cost for the new error rates (e.g., Field et al., 2005).

Failure to conduct planning studies and just 'do what you can with available resources' generally wastes resources on fruitless effort (Legg and Nagy, 2006; Taylor et al., 2007), provides a false sense of management performance (Fairweather, 1991), and perhaps most importantly, encourages continued institutional ignorance of the true costs of information. Planning results can help garner institutional support by making clear the length of commitment that may be required before achieving the desired information goals, e.g., 31 years in the case of monitoring brown bears at Togiak NWR. This can be especially important given the tendency for relatively rapid shifts of focus at the higher management levels. Planning results can also help in assessing the potential effects of declining support.

Study planning can be technically challenging and time consuming, but its cost is negligible compared to the potential savings from having identified the set of optimal designs and, thereby, also the ineffective or inefficient ones. The labor required to develop the current results totaled to less than a third of a year's focused labor by a senior biometrician (not counting interruptions by other duties). Unfortunately, a major barrier to regularly conducting such planning studies is a lack of availability of technical capacity within management agencies and/or their partners (Field et al., 2007). Natural resource managers should emphasize the importance of this need so that program staffing plans can take concrete steps to address it. Having completed just a few such studies is often sufficient to demonstrate their value.

Relevance of simulation results clearly depends on the adequacy of the process, observation, and estimation models (Seavy and Reynolds, 2007). It is unclear how the current results are affected by the simple process model (Table 1) that ignores any aspect of demographics, spatial patterns of habitat use, changing harvest rates or patterns, or changing system drivers such as climate or salmon-stock abundance. The modeled decline assumed that all age and sex classes declined at equal rates, rather than the least detectable individuals being most susceptible to loss or some other demographically inhomogeneous process. However, regardless of how desirable a more detailed process model may be, the model used here reflects the current level of detailed data available for modeling brown bear population dynamics in this region.

Observation errors likely could be reduced somewhat in actual practice by incorporating additional covariates, such as terrain topographic relief (e.g., high vs. low), into the detection function (e.g., Walsh et al., 2010) or combining information on detections across multiple monitoring surveys or surveys from similar habitats (e.g., Bayesian hierarchical modeling; Royle and Dorazio, 2008). Also, a potentially more precise survey estimator has been suggested (Laake and Borchers, 2004). Further, in real life, managers would avail themselves of information on population demographics collected during the line transect survey as well as information from other sources, such as harvests.

In actual application, effort should be made to remove the observation error from the estimated process variation, the latter being the relevant source of uncertainty and the former simply noise (Link and Nichols, 1994; de Valpine and Hastings, 2002; Lindley, 2003). Such an analysis could greatly improve monitoring effectiveness and efficiency, potentially greatly reducing the expected cumulative costs and the expected time to achievement of the management information objectives. However, that requires not only much more complicated analyses than employed here but rather strong insight into the most relevant process model.

The impossibility of synthesizing the simulation errors from the assumed process, observation, and estimation models (Table 1), reinforces the importance of revisiting the planning study as further surveys are conducted and system understanding improves (Field et al., 2007). Ideally, the planning study would be redone after each survey, leading to an adaptive monitoring program that continually learns and improves its efficiency and effectiveness (Lindenmayer and Likens, 2009). Having developed the planning tools, it is a relatively simple matter to investigate other scenarios of change as well as other levels of statistical confidence and power (Field et al., 2004, 2005).

5. Conclusion

A number of landscape-scale monitoring programs exist or are being initiated in the northern latitudes, among many others elsewhere, to aid and inform natural resource management in the face of a variety of sources of change. Based on past history, much of this effort risks being wasted (Legg and Nagy, 2006; Nichols and Williams, 2006; Field et al., 2007) unless the management agencies and their non-governmental partners embrace and engage in thorough program planning (Fancy et al., 2009; Lindenmayer and Likens, 2009, 2010) – from the formulation of management objectives and information needs through the numerous intervening stages to the selection of survey and analysis methods, culminating in the types of study design planning illustrated here. Achieving this will require expanding the technical capacity of most resource-management agencies and/or their non-governmental partners. In the long run, such costs should result in improved management efficacy and efficiency, earlier problem detection (and resolution), improved production efficiency, development of 'outside' clientele and program champions, greater institutional support, and improved institutional awareness of the true costs of information.

Acknowledgements

This work was funded by the US Fish and Wildlife Service Division of Realty and Natural Resources in Anchorage, Alaska, Togiak National Wildlife Refuge, and the National Park Service, Southwest Alaska Network Inventory and Monitoring Program. The ideas presented here were improved, often indirectly, through many conversations with Earl Becker, Patrick Walsh, Heather Renner, Emily Silverman, and Harold Laskowski. The manuscript benefited from the thoughtful comments of three anonymous reviewers.

Appendix A. Detailed observation model

We briefly describe the program developed to simulate all phases of the survey data collection and analysis method; the program was written for the R analysis environment (R Development Core Team, 2009) and is available from the first author. The findings and conclusions in this article are those of the authors and do not necessarily represent the views of the U.S. Fish and Wildlife Service.

A.1. Setting up the simulation 'region' and brown bear 'population'

The survey method requires that number and size of animal groups in the study region remain constant during the survey and their distribution remain constant relative to the time required to survey a transect (Becker and Quang, 2009). This was simulated by defining a square study region 145.53 km on a side, giving the same area as the Togiak NWR study (21,178 km^2). A specific brown bear density was simulated in the region by (i) calculating the total number of bears to place in the region to achieve that density, (ii) randomly simulating a large number of bear group sizes from the Togiak study group size distribution (Walsh et al., 2010), stopping when the cumulative number of bears first met or exceeded the total number required, (iii) randomly placing the bear groups in the square study region following a uniform spatial distribution. Bear group locations and sizes were held fixed for the duration of the simulations under that density.

A.2. Simulating observations

The detailed observation model explicitly simulated the dominant sources of uncertainty in the observation process: a sample of n 25-km-long transects were randomly selected in the study region and detection of each bear group along a transect was simulated for the 'pilot' and, independently, the 'observer' using their respective detection functions. These 'true' detection functions were simplifications of the functions identified in the Togiak NWR study (Table 5 of Walsh et al., 2010). The covariate model for the scale parameter was $\ln(\lambda) = \beta_0 + \beta_1 \ln(\mathrm{ESD})$, where ESD was the effective search distance, the furthest distance off the transect that pilot was actively searching immediately preceding a group's detection (see Becker and Quang (2009) for a discussion of this covariate). All parameters were set at the values from fitting these detection models to the Togiak study results. The ESD covariate was generated by adding a random number to the distance from the transect to the bear group. Investigation of the empirical distribution of these 'innovations' for the Togiak NWR survey, i.e., ESD_i – (distance from transect)$_i$ for each detected bear group i, revealed that they were adequately modeled by a gamma distribution (parameters shape = 1.27, rate = 0.004) (unpublished data), a selection originally considered because of its flexible shape and ability to represent skewed distributions. The simulations assumed that the size of a detected group was counted without error.

A.3. Simulating analysis

Estimation followed the method of Becker and Quang (2009), but without a model selection process – only the scale model above was fit during estimation. Standard errors were estimated from 600 bootstrap replicates for each simulated sample; this had been shown to be an adequate number of bootstrap replicates for precise standard error estimates (unpublished data). Analyses were conducted using the GammaMRDS library of functions (Reynolds et al., 2010) for the statistical package R (R Foundation, 2010).

Appendix B. Analytical approximation to detailed observation model

The detailed observation model was used to generate 500 estimated densities and their standard errors (using 600 bootstrap replicates) for each sample size n = {800, 1000, 1200, 1400, 1600} under a density of 40.3 bears/1000 km^2, and for n = 1000 transects under each density 20, 80, and 160 bears/1000 km^2. The resulting sampling distributions were used to identify analytical approximations for generating density estimates and their standard errors for any sample size and density (see Table 1).

B.1. Approximate sampling distribution for $\hat{D}$

Brown bear density estimates, $\hat{D}$, were approximately normally distributed around the true density (unpublished data). The standard deviation of this distribution was modeled by a two step process. First, for each simulated sample size and density, the variance of the 500 density estimates was calculated. The resulting values from different samples sizes under a common density were used to calculate the 'per unit sampling variance' for that density (unpublished data), a quantity allows one to easily estimate the sampling variance for that density under any sample size (described in Kish, 1965, p. 255). The per unit sample variances from different densities were found to be adequately modeled by a linear function of the true density (unpublished data), giving the formula for $\sigma^2_{\mathrm{Density}}(D_t, n)$ shown in Table 1.

B.2. Approximate sampling distribution for $\widehat{SE}(\hat{D})$

The natural logarithm of the estimated standard errors for the brown bear density estimates, $\log(\widehat{SE}(\hat{D}))$, were approximately normally distributed (unpublished data). Though an adequate approximation, the simulation results did include more extremely large values than this distribution. Thus the parameter values for each simulation scenario's normal distribution were fit to that scenario's 500 simulation results using a robust k-step estimator (Kohl, 2005) from the R library ROptEst (Kohl and Ruckdeschel, 2009), an estimation method that automatically downweights extreme observations. The mean of the normal distribution for any given sample size and density was adequately modeled as a linear function of the sample size and square-root of the density (Table 1) (unpublished data). The variance of the normal distribution for any given sample size and density was modeled following the same process above, with the modification that the per unit variance for a given density was found to be adequately modeled as a linear function of the logarithm of the density (given in Table 1) (unpublished data).

References

Becker, E.F., Quang, P.X., 2009. A gamma-shaped detection function for line-transect surveys with mark-recapture and covariate data. Journal of Agricultural, Biological, and Environmental Statistics 14, 207–223.

Beever, E.A., Woodward, A., in this issue. Design of ecoregional monitoring in protected areas of high-latitude ecosystems, under contemporary climate change. Biological Conservation.

Borchers, D.L., Buckland, S.T., Zucchini, W., 2002. Estimating Animal Abundance: Closed Populations. Springer-Verlag, London.

Boulanger, J., White, G.C., McLellan, B.N., Woods, J., Proctor, M., Himmer, S., 2002. A meta-analysis of grizzly bear DNA mark-recapture projects in British Columbia, Canada. Ursus (13), 137–152.

Boulanger, J., Proctor, M., Himmer, S., Stenhouse, G., Paetkau, D., Cranston, J., 2006. An empirical test of DNA mark-recapture sampling strategies for grizzly bears. Ursus (17), 149–158.

Fairweather, P.G., 1991. Statistical power and design requirements for environmental monitoring. Australian Journal of Marine and Freshwater Research 42, 555–567.

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.

Field, S.A., Tyre, A.J., Possingham, H.P., 2005. Optimizing allocation of monitoring effort under economic and observational constraints. Journal of Wildlife Management 69, 473–482.

Field, S.A., Tyre, A.J., Rhodes, J.M., Jonzen, N., Possingham, H.P., 2004. Minimizing the cost of environmental management decisions by optimizing statistical thresholds. Ecological Letters 7, 669–675.

Gende, S.M., Edwards, R.T., Willson, M.F., Wipfli, M.S., 2002. Pacific salmon in aquatic and terrestrial ecosystems. Bioscience 52, 917–928.

Gerrodette, T., 1987. A power analysis for detecting trends. Ecology 68, 1364–1372.

Gibbs, J.P., Snell, H.L., Causton, C.E., 1999. Effective monitoring for adaptive wildlife management: lessons from the Galápagos Islands. Journal of Wildlife Management 63, 1055–1065.

IUCN Standards and Petitions Subcommittee, 2010. Guidelines for Using the IUCN Red List Categories and Criteria, Version 8.0. <http://intranet.iucn.org/webfiles/docs/SSC/RedList/RedListGuidelines.pdf> (accessed 08.04.10).

Kish, L., 1965. Survey Sampling. Wiley and Sons, New York.

Kohl, M. 2005. Numrical Contributions to the Asymptotic Theory of Robustness. Dissertation, University of Bayreuth.

Kohl, M., Ruckdeschel, P., 2009. ROptEst: Optimally Robust Estimation. R Package Version 0.7. <http://robast.r-forge.r-project.org> (accessed 15.02.10).

Kovach, S.D., Collins, G.H., Hinkes, M.T., Denton, J.W., 2006. Reproduction and survival of brown bears in southwest Alaska, USA. Ursus 17, 16–29.

Laake, J.L., Borchers, D.L., 2004. Methods for incomplete detection at distance zero. In: Buckland, S.T., Anderson, D.R., Burnham, K.P., Laake, J.L., Borchers, D.L., Thomas, L. (Eds.), Advanced Distance Sampling: Estimating Abundance of Biological Populations. Oxford University Press, Oxford, pp. 108–190.

Larsen, D.P., Kincaid, T.M., Jacobs, S.E., Urquhart, N.S., 2001. Designs for evaluating local and regional scale trends. Bioscience 51, 1069–1078.

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.

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.

Lindenmayer, D.B., Likens, G.E., 2010. The science and application of ecological monitoring. Biological Conservation 143, 1317–1328.

Lindley, S.T., 2003. Estimation of population growth and extinction parameters from noisy data. Ecological Applications 13, 806–813.

Link, W.A., Nichols, J.D., 1994. On the importance of sampling variance to investigations of temporal variation in animal population size. Oikos 69, 539–544.

Lyons, J.E., Runge, M.C., Laskowski, H.P., Kendall, W.L., 2008. Monitoring in the context of structured decision-making and adaptive management. Journal of Wildlife Management 72, 1683–1692.

National Research Council, Committee on Management of Wolf and Bear Populations in Alaska, 1997. Wolves, Bears, and Their Prey in Alaska: Biological and Social Challenges in Wildlife Management. National Academy Press, Washington, DC.

Nichols, J.D., Williams, B.K., 2006. Monitoring for conservation. TRENDS in Ecology and Evolution 21, 668–673.

Olsen, A.R., Sedransk, J., Edwards, D., Gotway, C.A., Liggett, W., Rathbun, S., Reckhow, K.H., Young, L.J., 1999. Statistical issues for monitoring ecological and natural resources in the United States. Environmental Monitoring and Assessment 54, 1–45.

Olson, T.L., Putera, J.A., 2007. Refining Monitoring Protocols to Survey Brown Bear Populations in Katmai National Park and Preserve and Lake Clark National Park and Preserve. Alaska Region Natural Resources Technical Report NPS/AR/NRTR-2007-66, National Park Service, Anchorage, Alaska. <http://science.nature.nps.gov/im/units/swan/Libraries/Reports/OlsonT_2007_PMIS45148_Final_Report_AKRO_Report_Series_FINAL.pdf>.

R Development Core Team, 2009. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. <http://www.r-project.org/>.

Reid, L.M., 2001. The epidemiology of monitoring. Journal of the American Water Resources Association 37, 815–820.

Renner, H., Reynolds, J.H., Sims, M., Renner, M., 2010. Evaluating the power of surface attendance counts to detect long-term trends in populations of crevice-nesting auklets. Environmental Monitoring and Assessment. doi:10.1007/s10661-010-1664-4.

Reynolds, J.H., Russell, B., Becker, E.F., Quang, P.X., 2010. GammaMRDS: An R Package for Fitting Gamma-shaped Detection Functions to Mark-recapture Line-transect Data. US Fish & Wildlife Service, Division of Realty and Natural Resources, Anchorage, Alaska.

Rhodes, J.R., Tyre, A.J., Jonzen, N., McAlpine, C.A., Possingham, H.P., 2006. Optimizing presence–absence surveys for detecting population trends. Journal of Wildlife Management 70, 8–18.

Royle, J.A., Dorazio, R.M., 2008. Hierarchical Modeling and Inference in Ecology: The Analysis of Data from Populations, Metapopulations and Communities. Academic Press-Elsevier, San Diego.

Seavy, N.E., Reynolds, M.H., 2007. Is statistical power to detect trends a good assessment of population monitoring? Biological Conservation 140, 187–191.

Shrader-Frechette, K.S., McCoy, E.D., 1992. Statistics, costs and rationality in ecological inference. Trends in Ecology and Evolution 7, 96–99.

Sims, M., Bjorkland, R., Mason, P., Crowder, L.B., 2008. Statistical power and sea turtle nesting beach surveys: how long and when? Biological Conservation 141, 2921–2931.

Sims, M., Wanless, S., Harris, M.P., Mitchell, P.I., Elston, D.A., 2006. Evaluating the power of monitoring plot designs for detecting long-term trends in the numbers of common guillemots. Journal of Applied Ecology 43, 537–546.

Skalski, J.R., Ryding, K.E., Millspaugh, J.J., 2005. Wildlife Demography: Analysis of Sex, Age and Count Data. Elsevier Inc., Burlington, MA.

Taylor, B.L., Gerrodette, T., 1993. The uses of statistical power in conservation biology: the vaquita and the northern spotted owl. Conservation Biology 7, 489–500.

Taylor, B.L., Martinez, M., Gerrodette, T., Barlow, J., 2007. Lessons from monitoring trends in abundance of marine mammals. Marine Mammal Science 23, 157–175.

Thomas, L., Krebs, C.J., 1997. A review of statistical power analysis software. Bulletin of the Ecological Society of America 78, 126–139.

Thompson, W.L., Miller, A.E., Mortenson, D.C., Woodward, A., in this issue. Developing effective sampling designs for monitoring natural resources in Alaskan national parks: an example using simulations and vegetation data. Biological Conservation.

de Valpine, P., Hastings, A., 2002. Fitting population models incorporating process noise and observation error. Ecological Monographs 72, 57–76.

Walsh, P.W., Reynolds, J.H., Collins, G., Russell, B., Winfree, M., Denton, J., 2010. Application of a double-observer aerial line-transect method to estimate brown bear population density in southwestern Alaska. Journal of Fish and Wildlife Management 1, 47–58.