Ecological Applications, 18(8), 2008, pp. 1932–1955

MULTIPLE HYPOTHESIS TESTING AND THE DECLINING-POPULATION PARADIGM IN STELLER SEA LIONS

NICHOLAS WOLF[1,2,4] AND MARC MANGEL[1,2,3]

[1]*MRAG Americas, 303 Potrero Street, 42-201, Santa Cruz, California 95062 USA*
[2]*Center for Stock Assessment Research, 1156 High Street, University of California, Santa Cruz, California 95064 USA*
[3]*Department of Applied Math and Statistics, 1156 High Street, University of California, Santa Cruz, California 95064 USA*

Abstract. We describe a novel spatially and temporally detailed approach for determining the cause or causes of a population decline, using the western Alaskan population of Steller sea lions (*Eumetopias jubatus*) as an example. Existing methods are mostly based on regression, which limits their utility when there are multiple hypotheses to consider and the data are sparse and noisy. Our likelihood-based approach is unbiased with regard to sample size, and its posterior probability landscape allows for the separate consideration of magnitude and certainty for multiple factors simultaneously. As applied to Steller sea lions, the approach uses a stochastic population model in which the vital rates (fecundity, pup survival, non-pup survival) at a particular rookery in each year are functions of one or more local conditions (total prey availability, species composition of available prey, fisheries activity, predation risk indices). Three vital rates and four scaling functions produce twelve nonexclusive hypotheses, of which we considered 10; we assumed a priori that fecundity would not be affected by fishery activities or predation. The likelihood of all the rookery- and year-specific census data was calculated by averaging across sample paths, using backward iteration and a beta-binomial structure for observation error. We computed the joint maximum likelihood estimates (MLE) of parameters associated with each hypothesis and constructed marginal likelihood curves to examine the support for each effect. We found strong support for a positive effect of total prey availability on pup recruitment, negative effects of prey species composition (pollock fraction) on fecundity and pup survival, and a positive effect of harbor seal density (our inverse proxy for predation risk) on non-pup survival. These results suggest a natural framework for adaptive management; for example, the areas around some of the rookeries could be designated as experimental zones where fishery quotas are contingent upon the results of pre-fishing season survey trawls. We contrast our results with those of previous studies, demonstrating the importance of testing multiple hypotheses simultaneously and quantitatively when investigating the causes of a population decline.

Key words: *aerial survey; beta-binomial observation error; declining-population paradigm;* Eumetopias jubatus; *killer whales; multiple-hypothesis testing; optimal foraging; Steller sea lions; stochastic dynamic programming; walleye pollock*.

INTRODUCTION

The problem of determining the reason or reasons for a population decline, known as the declining-population paradigm (Caughley 1994), has traditionally been handled by testing one hypothesis at a time, often with an empirical (nonmechanistic) model. Common methods include testing for spatial or temporal correlation between population trends and an environmental index (Green 1995), running a simulation to determine whether a particular factor could have caused the decline (e.g., Barrett-Lennard et al. 1995), and fitting a matrix model to determine which transition was most likely impacted (e.g., York 1994).

Regardless of the method used, there is a problem that occurs when single-hypothesis approaches are applied to declining populations: Different studies may find seemingly definitive statistical support for different hypotheses. The problem is that virtually any hypothesis that predicts a drop in vital rates will appear to fit at least some of the data when the population is in decline. Different hypotheses are unlikely to have mutually exclusive predictions when they both predict a population decline: All of them ultimately translate into impacts on survival or fecundity. It is difficult to compare the support for different individual hypotheses without putting them together in the same model, because the fit of each model would change if alternative models were allowed to explain some of the same variation.

A second issue with many existing approaches is that their analyses are fundamentally empirical/statistical, testing for correlation without specifying an underlying

Manuscript received 31 July 2007; accepted 8 January 2008; final version received April 2008. Corresponding Editor: S. S. Heppell.
[4] E-mail: nwolf@alumni.sfu.ca

behavioral, ecological, or evolutionary mechanism. The resulting functions can only be trusted within the range of conditions to which they were fitted. If the population then moves outside of that range (as declining populations continuously do), the shape of the extrapolated function may not be appropriate (Norris 2004). In order to minimize this problem, it is important to formulate hypotheses mechanistically whenever possible.

Thus, the declining-population paradigm is in need of a unified approach that emphasizes mechanism and theory. The approach must allow for multiple hypotheses to be tested simultaneously, using a common currency (likelihood), so that a joint solution can be found for the strengths of all proposed effects. This concept of confronting multiple models with the data is classic ecological detection (Hilborn and Mangel 1997). Constructing the model as a likelihood framework makes it possible to fit the parameters using all available spatial and temporal variation, including incomplete datasets. Formulating the hypotheses as mechanistic functions makes it possible to interpret the results directly and extrapolate beyond the observed conditions. This approach offers the distinct advantage of quantifying both the relative contributions of different drivers, and our uncertainty about them, making it a valuable tool for decision-making in management of declining and threatened species.

Test case: the Steller sea lion decline

The population of Steller sea lions (*Eumetopias jubatus*) in western Alaska has fallen by >80% since the late 1970s (Fig. 1). The initial decline was characterized by a loss of ~15% of the population per year, accompanied by reduced size-at-age and other symptoms of nutritional stress (Castellini 1993, Calkins et al. 1998). After 1989 or so, the rate of decline slowed to ~5% per year (Sease and Loughlin 1999). The available evidence suggests that animals in the declining western population were in better condition than those in the growing eastern population at this time (Andrews et al. 2002), in line with the general expectation for a population reduced far below carrying capacity (e.g., Laidre et al. 2006). The western stock was listed as endangered in 1997.

Despite the efforts of many scientists to identify a "smoking gun" responsible for the decline, there is still no definitive answer (for a thorough review, see NRC 2003). However, there is growing consensus that the initial decline involved nutritional stress (Trites and Donnelly 2003) and low juvenile survival (York 1994, Holmes and York 2003, Holmes et al. 2007), whereas the later decline (after about 1990) may have resulted from a decline in fecundity (Holmes and York 2003, Holmes et al. 2007). There is also a correlation between population growth rate and diet diversity among rookeries, suggesting that the availability of multiple prey types provides a nutritional benefit or minimizes the impact of a shortfall in any single prey type (Merrick et al. 1997).

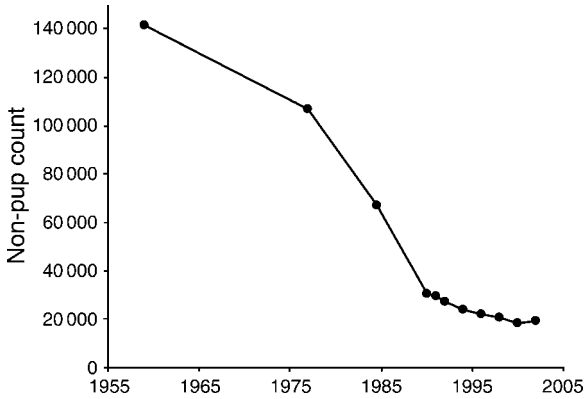


FIG. 1. Total non-pup counts for the western population of Steller sea lions (*Eumetopias jubatus*), ca. 1959–2002, showing the 11 years when range-wide totals were available (Merrick et al. 1987, National Marine Fisheries Service). The western population extends from 144° W longitude to the western tip of the Aleutian island chain. The sheer magnitude of the decline is obvious in this composite figure, but the spatial structure is not.

The Steller sea lion (SSL) decline in western Alaska was preceded by declines in the populations of northern fur seals, *Callorhinus ursinus*, and Pacific harbor seals, *Phoca vitulina*, occupying the same region. The causes of these declines remain similarly unexplained (Merrick 1997). One possibility is that killer whales, *Orcinus orca*, have depleted the most profitable prey types over time and progressively expanded their diet to include less and less profitable prey. This "diet breadth" hypothesis has received considerable attention recently (Springer et al. 2003, Wade et al. 2003, DeMaster et al. 2006, Mangel and Wolf 2006).

The numerous hypotheses that have been proposed to explain the decline fall into five main categories: (1) insufficient prey availability, (2) unsuitable prey species composition, (3) direct mortality related to fishing activities, (4) enhanced depredation by killer whales or sharks, and (5) disease or pathogen outbreaks. There is some evidence in support of each of these, but none is conclusive (Ferrero and Fritz 2002, NRC 2003).

As explained above, the lack of consensus results in part from the fact that the different hypotheses make nonexclusive predictions regarding the population trends to which they are fitted. A second issue is that the historical data (rookery counts based on aerial photos taken during the breeding season) are both noisy and sparse, leading most researchers to sacrifice some useful variation by pooling data across rookeries or across years (compare Figs. 1 and 2). This exacerbates the first problem, because the spatial and temporal pattern of the decline is crucial for differentiating between alternative effects. Thus, different researchers have found support for different hypotheses, leading to confusion and dissension.

The recent NRC (2003) report asserts that "finer-scale spatial analysis of Steller sea lion populations and

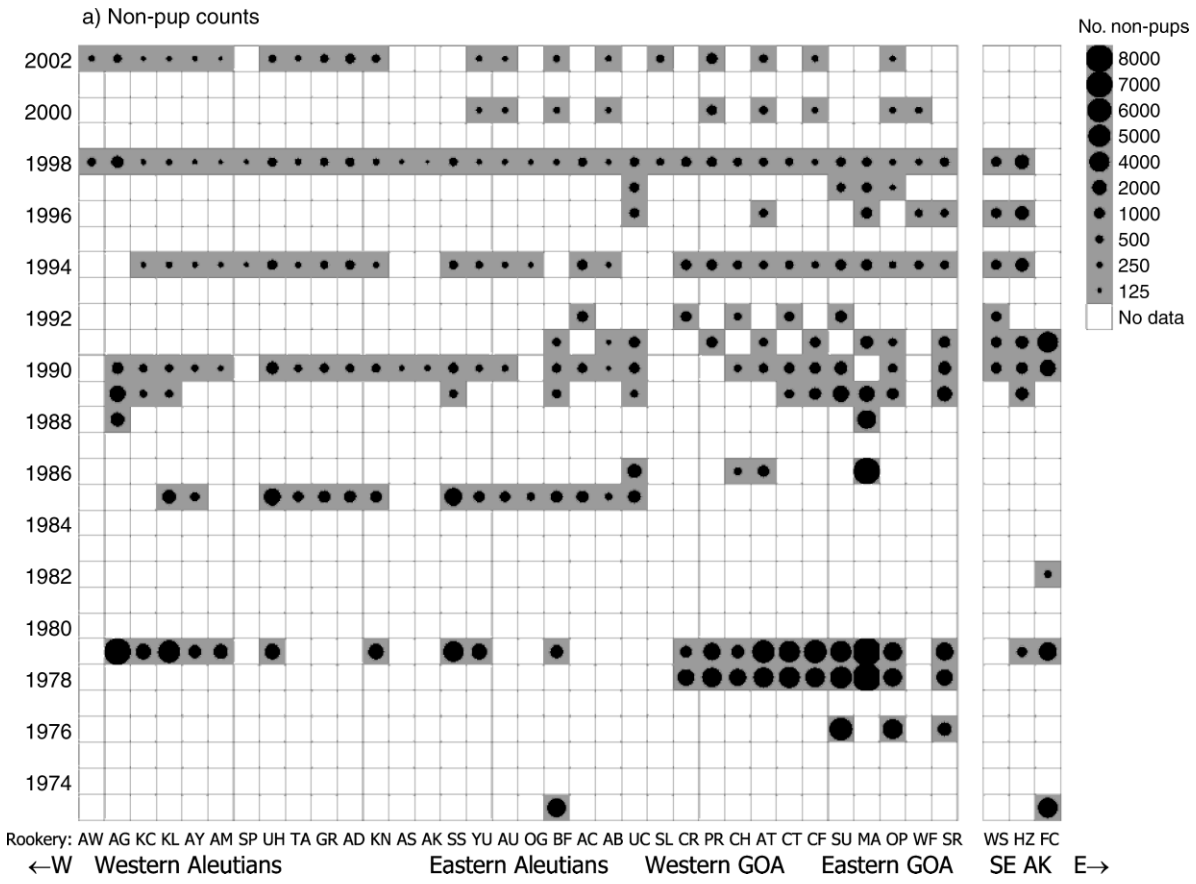


FIG. 2. Steller sea lion counts for individual rookeries in the Aleutian Islands, (GOA, Gulf of Alaska), and southeast Alaska (SE AK), 1973–2002. Even though not every population was censused in every year, many more data are available at this scale. Each column represents a rookery (see Appendix A: Table A1 for full rookery names); each row is a year. Panels (a) and (b) show non-pup and pup counts, respectively. Panel (c), in which the pup count has been divided by the non-pup count, provides an index of fecundity.

environmental conditions will be required to uncover potential region-specific determinants that are affecting sea lion survival" and calls for new modeling approaches that are both spatially and temporally explicit. Our work is aimed squarely at this gap.

The only previous studies of this system that considered multiple hypotheses simultaneously did so using Ecopath/Ecosim models and spatially pooled data (NRC 2003, Guenette et al. 2006). The NRC (2003) model was not able to predict a decline as steep as that observed, but the Guenette et al. (2006) model had better success and found qualitative support for all four factors considered (predation, ocean productivity, fisheries, and competition with other species). Our study goes several steps further, using all available spatial and temporal variation and analyzing the data using a mechanistic likelihood-based model.

Steller sea lion details

Steller sea lions occupy the north Pacific coastline from central California to Japan. Evidence from mitochondrial and nuclear DNA shows that the population east of 144° W longitude is genetically distinct from the population west of that line (Bickham et al. 1998, Hoffman et al. 2006), and an analysis of resightings of marked individuals confirms that there is very little exchange of breeding individuals between the two regions. Females tend to return either to their natal rookery (~67% of the time in the western stock) or to a neighboring rookery (Raum-Suryan et al. 2002).

Rookeries are generally located on small, remote islands (for rookery locations see Appendix A: Table A1). Most pups (one per pregnant female) are born within a two-month period centered in June (Pitcher et al. 2002), and enter the water for the first time when they are 2–4 weeks old (Sandegren 1970). Mothers alternate between nursing on land and foraging at sea, leaving the pups behind to fend for themselves for 1–2 days at a time (NRC 2003). Pups depend on their mothers for nourishment throughout most the first year, during which time they gradually learn to forage and become proficient at diving. Maximum dive depth increases from

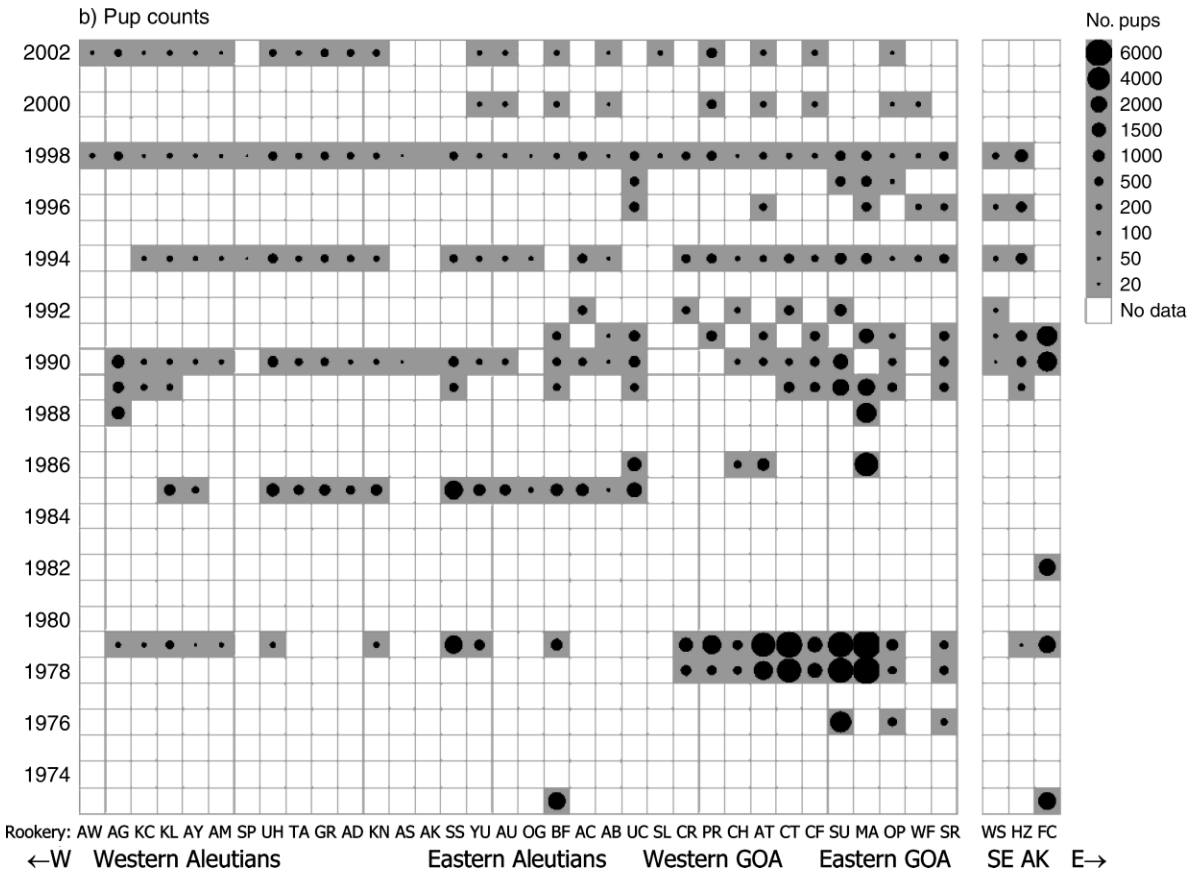


FIG. 2. Continued.

70 m at 6 months of age to 140 m at 12 months (Rehberg et al. 2001) and 250 m in adults (Merrick and Loughlin 1997).

Steller sea lions are largely opportunistic foragers keying in on locally and temporally aggregated prey (including walleye pollock, herring, eulachon, salmon, and Pacific cod), although they may display some prey preference under certain conditions. Walleye pollock (*Theragra chalcogramma*) are currently the principal diet component for both the western and eastern stocks (Alverson 1992, Anderson and Blackburn 2002). Other important prey species include Atka mackerel (*Pleurogrammus monopterygius*), Pacific cod (*Gadus macrocephalus*), and Pacific herring (*Clupea pallasi*). After they wean, but before they become proficient at deep diving and catching fast-swimming fish, pups may rely heavily upon slow-moving, easily captured prey, including shrimp (Hansen 1997). The abundance of such prey species could be important in determining the recruitment success of pups (Merrick and Loughlin 1997).

Population estimates of Steller sea lions at each rookery are obtained from counts based on aerial photos taken during the breeding season. Pups are counted by observers on the ground (Holmes and York 2003).

METHODS

Our general approach was to confront alternative models with the data and sort them according to their ability to reproduce it (Hilborn and Mangel 1997, Burnham and Anderson 2002). That is, we did not subscribe to a particular hypothesis and set out to evaluate our idea in isolation. Rather, we sought to understand the role of multiple mechanisms in the decline of the Steller sea lion. The analysis can be described in five general steps:

1) Find estimates of maximum (pre-decline) vital rates and construct a simple population model for the study species, including process stochasticity. The pre-decline vital rates provide background or default values that are modified according to each hypothesis.

2) Identify all plausible hypotheses to explain the decline. Formulate each as a simple function that scales down one or more vital rates according to local conditions; this resolves the problem of a common currency for the different hypotheses. Ideally, each function should be mechanistic and minimally parameterized, with one end of the range of possible parameter values corresponding to "no effect."

3) Compile matrices of local-conditions data sorted to the same levels of spatial and temporal detail available

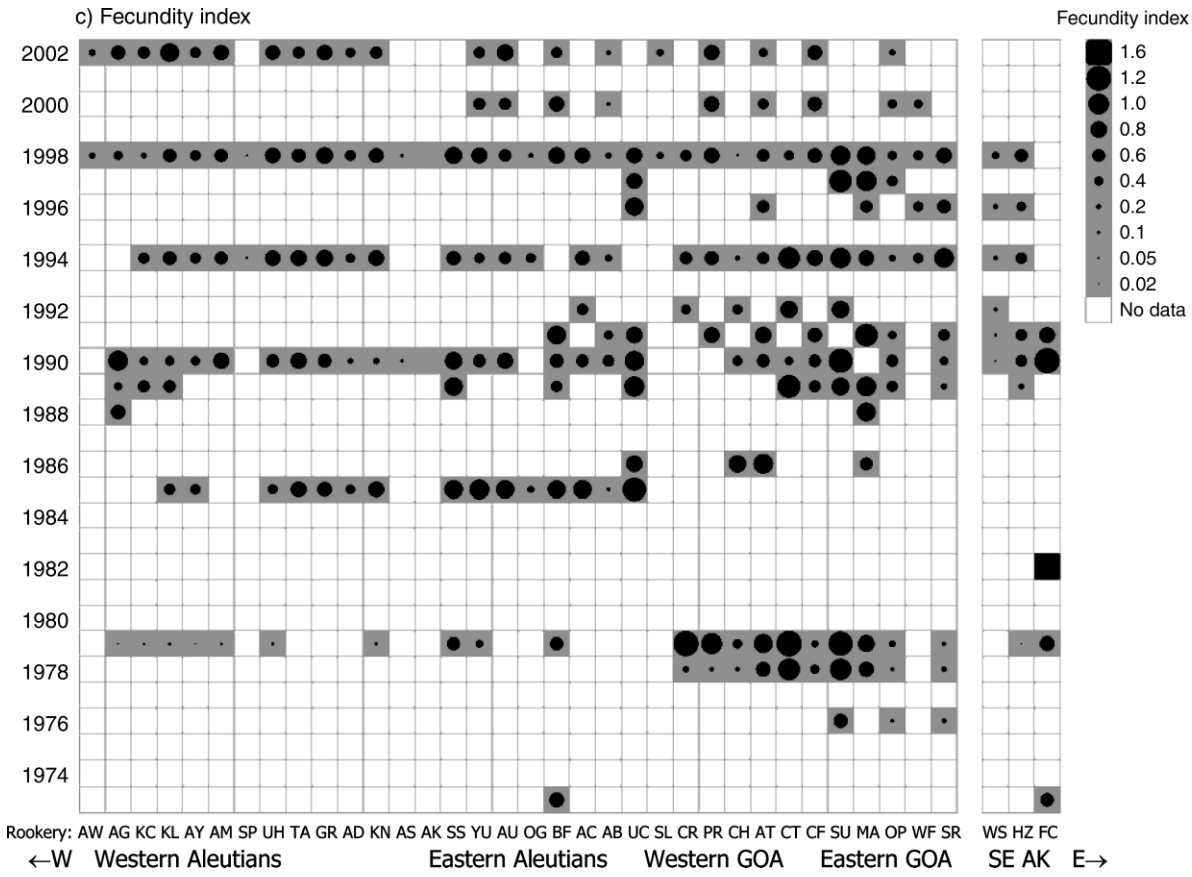


FIG. 2. Continued.

in the population data. These data provide the model inputs that modify the vital rates locally.

4) Characterize the observation error distribution, typically by fitting a binomial, Poisson, or beta-binomial distribution to a set of replicate counts. This allows for the calculation of the likelihood of observations.

5) Simultaneously estimate the parameters in all vital rate functions by searching the multi-dimensional parameter space, calculating the likelihood of all observed population data for each combination of values, and finding the joint maximum-likelihood solution. Find marginal likelihoods for each parameter and estimate confidence intervals around the maximum-likelihood estimates (note that marginal likelihood distributions are proportional to posterior probability distributions when uniform priors are used). Interpret the confidence intervals in terms of the range of strength of the effect. Any hypotheses for which the confidence interval only includes values where the function has negligible effect are considered unsupported by the data. All other hypotheses are supported.

An important point is that the omission of a major factor as a hypothesis in the model might not prevent it from appearing to fit the data reasonably well, but the results will be distorted. This caveat applies to all models, but our approach made it easier to understand the mechanism behind it: The estimated parameter values of the included factors will shift to explain some of the variation that would have been explained by the missing factor. At the extreme, a hypothesis that explains a significant amount of variation when considered alone may have no effect when allowed to compete with other hypotheses. Our model included all plausible hypotheses for which the available data were of sufficient temporal and spatial extent and detail to support the analysis.

The Steller sea lion population model: two age classes, three vital rates

We assumed that the population dynamics can be satisfactorily described in terms of the two age classes readily available from the rookery counts: pups (age 0–1) and non-pups. While it is possible to obtain a rough estimate of the juvenile fraction by comparing the lengths of animals with that of an adult male in the same aerial photo (Holmes and York 2003), these data are not available for all rookeries, and we did not include a juvenile age class in our model. We followed York (1994) in considering only females and assumed a

50:50 sex ratio (NRC 2003). Thus, the underlying variables are:

$$\begin{aligned} J(i,t) &= \text{number of female pups at rookery } i \text{ in year } t. \\ N_{\text{true}}(i,t) &= \text{number of female non-pups at rookery } i \text{ in year } t. \end{aligned} \quad (1)$$

We assumed that the number of pups observed equals $J(i, t)$, but that the number of non-pups observed, $N_{\text{obs}}(i, t)$, is less than or equal to the actual number of non-pups, $N_{\text{true}}(i, t)$, because some unknown fraction of non-pups are foraging at sea at the time of the census. The best available distribution for this type of observational uncertainty is the beta-binomial (Martz and Waller 1982, Evans et al. 2000; see Appendix B), with parameters α and β:

$$\Pr\{N_{\text{obs}} = k \mid \alpha, \beta, N_{\text{true}}\} = \binom{N_{\text{true}}}{k} \frac{\Gamma(\alpha + k)\Gamma(\beta + N_{\text{true}} - k)\Gamma(\alpha + \beta)}{\Gamma(\alpha + \beta + N_{\text{true}})\Gamma(\alpha)\Gamma(\beta)} \quad (2)$$

where $\Gamma(x)$ is the gamma function (Abramowitz and Stegun 1964) satisfying the recursion $\Gamma(x + 1) = x\Gamma(x)$.

Using likelihood estimation based on eight replicate counts from a single rookery (Outer/Pye Island) in 1992, we found the maximum likelihood estimates (MLEs) $\alpha = 6$ and $\beta = 2$ (see Appendix B).

Because the breeding season is relatively compressed and the population size is small, we assume discrete time dynamics and no density dependence. In that case, there are three fundamental parameters at rookery i in year t:

$$\begin{aligned} \rho_{i,t} &= \text{probability of pup recruitment (survival from birth to non-pup status at age 1)} \\ \sigma_{i,t} &= \text{probability of non-pup survival from year } t \text{ to year } t + 1 \\ \phi_{i,t} &= \text{per capita probability for non-pups of successful reproduction in year } t. \end{aligned} \quad (3)$$

We set fixed background values for these parameters (denoted by ρ_0, σ_0, and ϕ_0) that are modified by local conditions according to parameterized functions reflecting each hypothesis. We estimated the background values from life tables based on data collected on the Marmot Island rookery (Calkins and Pitcher 1982) and used by Pascual and Adkison (1994) and York (1994), with corrections in Holmes and York (2003). The annual growth rate of a population using the original life table was 1.09%. Because our model pools all age classes above one year, we calculated σ_0 and ϕ_0 by integrating the survival and fecundity probabilities across all non-pup classes in the stable age distribution produced by the life table and found the following:

$$\rho_0 = 0.776 \qquad \sigma_0 = 0.858 \qquad \phi_0 = 0.197. \quad (4)$$

This fecundity estimate accounts for ~50% of the pups being female (York 1994, NRC 2003) and for ~45% of the non-pup population being juvenile (Holmes and York 2003). We also scaled the non-pup counts by a factor of 0.75 to account for the observation that roughly 75% of the non-pups on rookeries are female (e.g., Trites and Larkin 1996, Gerber and Van Blaricom 2001).

We used Steller sea lion counts available in the online database of the Alaska Fisheries Science Center (AFSC)/National Marine Mammal Laboratory (*available online*).[5] Table A3 in Appendix A lists all other parameters and sources of data.

All analyses were performed using MATLAB software, version 7 (Mathworks 2004).

Hypotheses considered

Of the numerous hypotheses that have been proposed to explain the decline of the western population (Ferrero and Fritz 2002), we considered a subset for which sufficient data are available to apply appropriate analytical and statistical methods (Table 1).

H_1–H_3: hypotheses related to prey quantity

According to hypotheses H_1 through H_3, low local abundance of groundfish and other prey causes local fecundity (H_1), pup recruitment (H_2), or non-pup survival probability (H_3) to be diminished. Specifically, we assumed that termination of pregnancy (H_1) or else starvation (H_2, H_3) occurs if a sea lion fails to find prey for a certain number of days, which we will call the maximum shortfall period, S_i. We define "local abundance" as the total estimated density of all prey within a 300-km swimming radius (r_{forage}; see details in eighth paragraph of this section) from the rookery. The 300-km cutoff, which was first used by Gerber and Van Blaricom (2001), is intended to encompass the full foraging range of a Steller sea lion.

Because very little solid information is available regarding the lengths of the shortfall periods corresponding to each vital rate, we used the rough figures of $S_1 = 14$ days, $S_2 = 7$ days, and $S_3 = 21$ days in the calculations for H_1, H_2, and H_3, respectively. The one-week maximum fasting period for recruiting pups and the three-week period for non-pups were chosen because they are slightly longer than typical fasting durations observed in the field (Rea et al. 2000, Milette and Trites 2003; D. Noren, *personal communication*) or imposed by researchers on captive animals (Rosen and Trites 2002, Trites and Porter 2002). The two-week period for termination of pregnancy falls in between the other two, reflecting the assumptions that (1) a pregnant female has more energy stores than a pup, and (2) she would sooner terminate her pregnancy than starve. Intake rate certainly affects body condition, and body

[5] ⟨http://nmml.afsc.noaa.gov/AlaskaEcosystems/sslhome/stellerhome.html⟩

TABLE 1. The 10 hypotheses and their respective parameters.

Hypothesis	Parameter and mechanism of decline
Category I: insufficient prey availability	
H_1: lower prey density → lower fecundity rate	c_1, prey density increment per unit encounter rate; foraging shortfalls terminate pregnancy
H_2: lower prey density → less pup recruitment	c_2, prey density increment per unit encounter rate; foraging shortfalls cause pup starvation
H_3: lower prey density → lower non-pup survival	c_3, prey density increment per unit encounter rate; shortfalls cause non-pup starvation
Category II: unsuitable prey species composition	
H_4: higher pollock fraction → lower fecundity rate	c_4, exponent of non-pollock prey fraction in fecundity multiplier
H_5: higher pollock fraction → less pup recruitment	c_5, exponent of non-pollock prey fraction in pup recruitment multiplier
H_6: higher pollock fraction → lower non-pup survival	c_6, exponent of non-pollock prey fraction in non-pup survival multiplier
Category III: direct mortality due to fishing activities	
H_7: more fishing activity → less pup recruitment	c_7, pup mortality rate per fishery gear deployment within 20 km of rookery
H_8: more fishing activity → lower non-pup survival	c_8, non-pup mortality rate per fishery gear deployment within 20 km of rookery
Category IV: enhanced depredation by killer whales or sharks	
H_9: fewer harbor seals (more predation) → less pup recruitment	c_9, fraction of potential pup recruitment lost when harbor seal density $< h_{crit}$
H_{10}: fewer harbor seals (more predation) → lower non-pup survival	c_{10}, fraction of potential non-pup survival lost when harbor seal density $< h_{crit}$
H_9, H_{10}	h_{crit}, harbor seal density below which sea lions become prey to killer whales

condition is known to affect the birth rate of pregnant Steller sea lions (Pitcher et al. 1998).

We sought a scaling function, ω_n, by which to multiply each "background" or pre-decline vital rate to account for each hypothesis ($n = 1, 2, 3$) and the local conditions. Our approach was first to calculate the daily probability that a sea lion fails to find food, and then to calculate the probability that an animal manages to survive through "winter" (a period of 180 days) without experiencing a sequence of consecutive unsuccessful foraging days exceeding its maximum fasting period. For the first part, we assumed that prey encounters are Poisson-distributed (Mangel 2006), and the daily probability of zero prey encounters at rookery i during year t, $\Pr_n\{\text{no prey} \mid i, t\}$, is therefore

$$\Pr_n\{\text{no prey} \mid i, t\} = e^{-\lambda_{\text{prey}}(i,t)/c_n} \tag{5}$$

where $\lambda_{\text{prey}}(i, t)$ is prey density expressed as total prey biomass (excluding pollock; see subsection *H_4–H_6: hypotheses related to prey quality*) recovered per unit volume of standardized trawling effort within the foraging area around rookery i during year t, and c_n is a fitted parameter for hypothesis n representing the increment in prey density required to increase the daily encounter rate by a unit amount.

For the second part, we used backward induction (Mangel and Clark 1988, Mangel 2006) to calculate the probability of surviving winter without experiencing any fatal shortfalls. This iterative calculation starts at the end of the season and moves backwards through time, effectively considering all possible combinations of 180 successful or unsuccessful foraging days. Given any date d and hunger state x (days since last meal), the probability of avoiding fatal shortfalls from day t to the end of the season when hypothesis n is operative, $F_n(x, d, i, t)$, satisfies the equation of stochastic dynamic programming:

$$F_n(x,d,i,t) = (1 - \Pr_n\{\text{no prey} \mid i,t\})F_n(0,d+1,i,t) + \Pr_n\{\text{no prey} \mid i,t\}\begin{cases} F_n(x+1,d+1,i,t) & \text{if } x+1 \le S_n \\ 0 & \text{if } x+1 > S_n. \end{cases} \tag{6}$$

At the last iteration, we arrive at $F_n(0, 1, i, t)$, which is the probability of avoiding starvation from the beginning to the end of the season. This is the scaling factor we used to reduce the background fecundity or survival rates according to local conditions and hypotheses H_1 through H_3:

$$\omega_n(i,t) = F_n(0,1,i,t). \tag{7}$$

The scaling factors ω_1, ω_2, and ω_3 correspond to hypotheses H_1, H_2, and H_3, respectively. The function ω_n is similar to a Type III functional response (Holling 1959) with respect to local prey density, $\lambda_{\text{prey}}(i, t)$ (Fig. 3a).

Note that if $c_n = 0$, then $\Pr_n\{\text{no prey} \mid i, t\} = 0$ and $\omega_n(i, t) = 1$, so that the background vital rate is multiplied by 1 and hypothesis H_n has no effect. Our

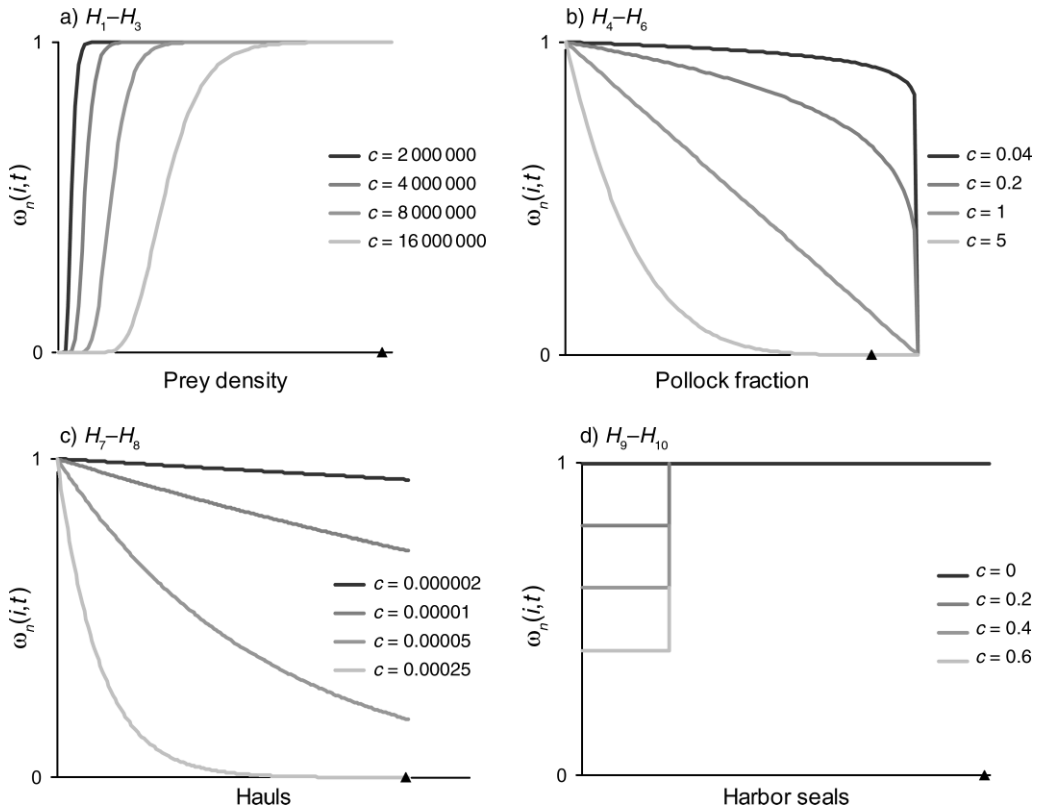


FIG. 3. Functional forms of the scaling factor $\omega_n(i, t)$ corresponding to each hypothesis ($n = 1, 2, \ldots, 10$). The probability of fecundity, pup recruitment, or non-pup survival at a particular rookery (i) in a particular year (t) is calculated by scaling down the maximum/background rate by the appropriate $\omega_n(i, t)$ factors, according to the local conditions. Different values of the fitted parameter c_n lead to different curves. The solid triangle in each panel indicates the maximum observed value of the environmental condition. (a) Prey quantity hypotheses (H_1– H_3); the x-axis shows total prey biomass trawled per unit effort (kg/km^3), excluding pollock. (b) Prey quality hypotheses (H_4–H_6); the x-axis is pollock catch as a fraction of total prey biomass. (c) Fisheries interactions (H_7–H_8); the x-axis is commercial fishing activity (number of hauls) within 20 km of the rookery. (d) Harbor seal abundance as an inverse proxy for predation risk (H_9–H_{10}); the x-axis is harbor seal abundance within 300 km of the rookery.

ultimate goal was to determine both the strength and statistical support for each hypothesis by finding the MLE of the c_n parameter and examining its marginal likelihood distribution, respectively.

We estimated local prey density for each rookery/year combination from the National Marine Fisheries Service's (NMFS) groundfish survey data by calculating the total catch per unit effort (CPUE) of all relevant prey species in trawls conducted within 300 km of the rookery in that year (Fig. 4a). Areas within 300 km of a rookery but requiring more than a 300-km swim to reach (because of an intervening land mass) were excluded. Data from the 12 months leading up to a breeding season (1 July–30 June) were relevant to the survival rates through that period and the fecundity rate at the end of it. In all but one year, the survey trawls were conducted in the intervals of July–October and May–June. (The exception is the 2000–2001 estimate, which includes trawl data from February and March.) We assumed that these data are representative of the entire 12-month period. There is certainly seasonal variability in the availability of prey, but the data are not sufficiently detailed for us to include this variation.

In the calculation of $\lambda_{\text{prey}}(i, t)$, we included 10 prey taxa that were described as "dominant," "important," or "most common" prey types (excluding walleye pollock) in a review of Steller sea lion diet studies in the 2001 SSL restricted areas Supplemental Environmental Impact Statement (*available online*; see also Sinclair and Zeppelin 2002).[6] The 10 prey types are listed in Appendix A: Table A2.

H_4–H_6: hypotheses related to prey quality

According to H_4, H_5, and H_6, fecundity, pup recruitment, and non-pup survival probability, respectively, are decreasing functions of the fraction of pollock in the environment. Specifically, starvation (H_5, H_6) or

[6] ⟨http://www.fakr.noaa.gov/sustainablefisheries/seis/sslpm/final/⟩

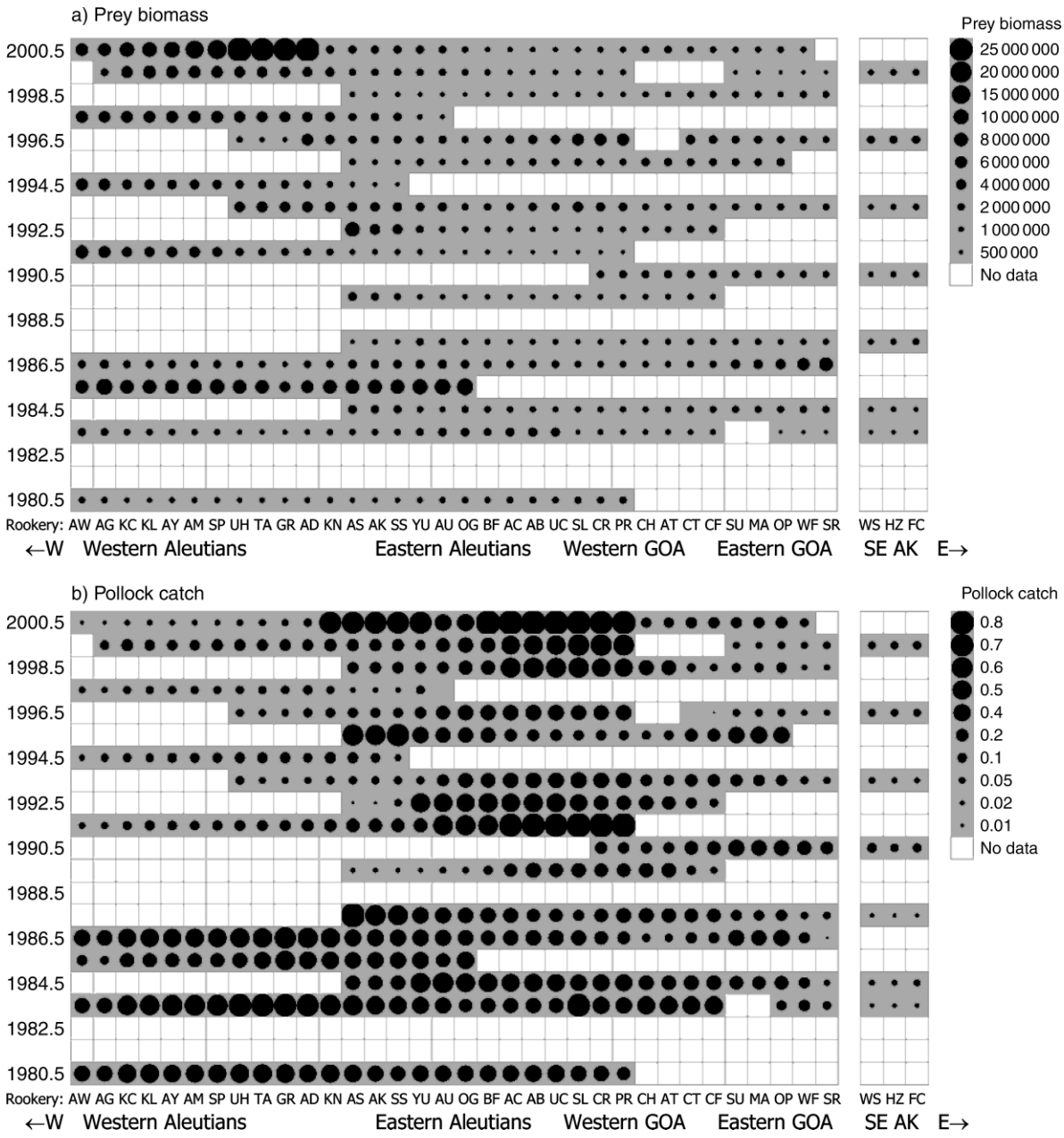


FIG. 4. Space–time plots for local conditions (within 300-km swimming distance of each rookery, unless otherwise specified). (a) Total prey biomass trawled per unit effort (kg/km^3), excluding pollock. (b) Pollock catch as a fraction of total prey biomass. (c) Commercial fishing activity (hauls) within 20 km. (d) Harbor seal abundance within 300 km. The year labels indicate that the data span pairs of years; for example, "2000.5" corresponds to the time period between mid-2000 and mid-2001. See Appendix A: Table A1 for full rookery names.

termination of pregnancy (H_4) occur with higher probability where prey other than pollock are relatively scarce.

These hypotheses are based on the observations that: (1) high fractions of pollock in the environment correlate with poor performance by Steller sea lions (e.g., Merrick et al. 1997, Rosen and Trites 2000), and (2) the pollock fraction has grown enormously since the mid-1970s. The exact mechanism for the negative impact is unclear. One possibility is that pollock are a poor food source, but the opportunistically foraging sea lions eat them anyway (Alverson 1992). Another possibility is that sea lions (particularly pups and juveniles) cannot (or will not) catch pollock, and they require other, more nutritious or more easily captured prey whose abundance scales inversely with the pollock fraction. However, hypotheses $H_1/H_2/H_3$ are designed to evaluate this possibility independently.

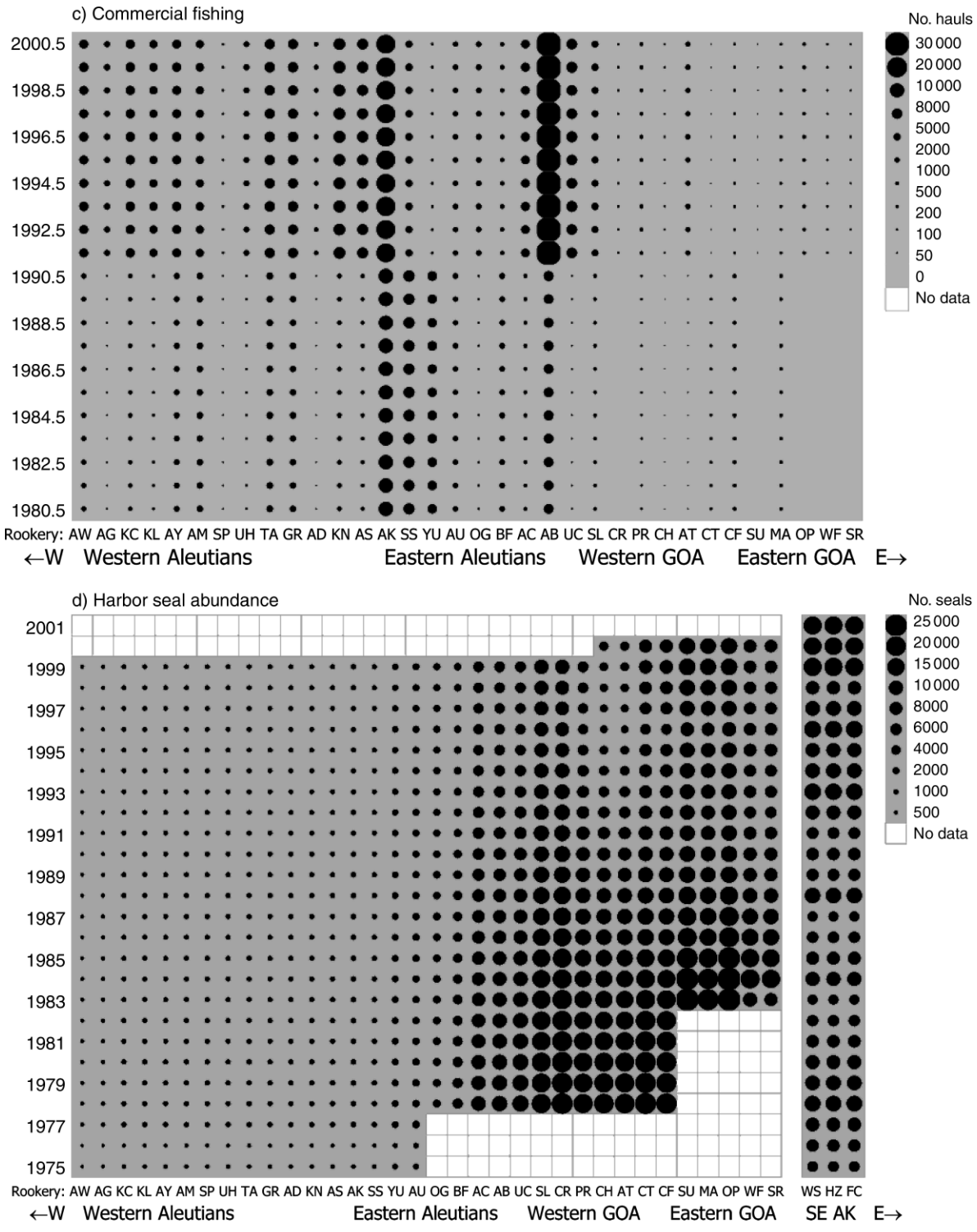


FIG. 4. Continued.

We model hypotheses 4–6 by multiplying the background vital rates by the fraction of non-pollock prey in the local environment (estimated from the NMFS groundfish survey trawl data), raised to the power, c_n (see Fig. 3b):

$$\omega_n(i,t) = \left[\frac{\lambda_{\text{prey}}(i,t)}{\lambda_{\text{prey}}(i,t) + \lambda_{\text{pollock}}(i,t)}\right]^{c_n}. \quad (8)$$

If $c_n = 0$, then $\omega_n(i, t) = 1$, and the pollock fraction has no effect on the corresponding vital rate. If $c_n = 1$, the effect is linear between the two extremes, with the vital rate of interest unaffected at 0% pollock and completely suppressed at 100% pollock. Values greater than 1 indicate strong suppression of the vital rate whenever pollock represents a significant fraction of the total biomass in the environment.

We calculated local pollock fractions from the NMFS groundfish survey data described above (Fig. 4b). Note that our methods do not require or imply any particular mechanism for the shift toward a pollock-dominated ecosystem. We simply used the trawl data as an input in the model, without addressing the ongoing controversy regarding relative contributions of climate change and fishing activities to the shift.

H_7, H_8: hypotheses related to fishing activity

Hypotheses H_7 and H_8 posit that survival probability of pups (H_7) or non-pups (H_8) is a decreasing function of local fishing intensity. These hypotheses reflect the observation that lethal interactions with commercial fishing boats do occur and may be locally important (but we assumed a priori that fecundity would not be affected by fishery activities). Incidental mortality (typically entanglement in fishing gear) was recently estimated to be killing <100 animals per year (Perez and Loughlin 1991, Loughlin and York 2002), but at certain times and in certain places (e.g., the Shelikov Strait trawl fishery in the late 1970s and early 1980s), it was probably a much bigger problem (NRC 2003). Deliberate shooting became illegal after 1990, but anecdotal reports suggest that it still occurs (NRC 2003). The magnitudes of these impacts are not well known, but they might be expected to scale with the number of fishing events (gear deployments) that occur near each rookery.

We need to estimate the magnitude (0–1) of the mortality rate c_n ($n = 7, 8$) per gear deployment within a certain distance of a rookery and compute the probability of not dying in $\mu(i, t)$ potential gear encounters. We therefore modeled Hypotheses 7 and 8 by reducing survival rates (from the background levels) according to the following function (Fig. 3c):

$$\omega_n(i, t) = (1 - c_n)^{\mu(i,t)}. \quad (9)$$

We used gear deployment data compiled by Daniel Hennen (Alaska Sea Life Center, Seward, Alaska, USA) from commercial groundfish fishery observer data provided by NMFS and observer coverage correction factors from a variety of sources. Using the same data, Hennen (2006) had found a strong negative correlation between the SSL population trend and the number of gear deployments within a 20-km radius ($r_{\text{fisheries}}$) around each rookery, for the period from 1977 to 1991. He had tried several spatial scales before determining that the 20-km radius provided the strongest effect; thus, we adopted the same radius. The data he provided were pooled across two time periods, 1977–1991 and 1991–2000, in order to avoid the problem of sparse coverage in individual years. All types of gear were counted, including pots, longline, and various types of trawl gear (Fig. 4c). See Hennen (2006) for details.

H_9, H_{10}: hypotheses related to predation mortality

According to the final two hypotheses, survival probability of sea lions declines when local harbor seal density falls below a critical threshold and their predators (particularly killer whales, *Orcinus orca*) expand their diet to include Steller sea lion pups (H_9) or non-pups (H_{10}).

Hypotheses 9 and 10 are motivated by optimal foraging theory (Stephens and Krebs 1986, Clark and Mangel 2000): When an optimal forager encounters a given prey type, it is predicted to accept the item only if its profitability meets or exceeds the expected long-term intake rate of a diet including all more profitable prey types. Profitability is defined here as the ratio of prey body mass to handling time, divided by the number of individuals sharing the prey. When the more profitable prey are scarce, less profitable prey are predicted to be included in the diet. In the case of killer whales, Steller sea lions are predicted to be included in the diet when the density of harbor seals (a more profitable prey type, according to estimates of body size and handling time) falls below a threshold (Mangel and Wolf 2006).

The range of the sea lions' western population is known to host a large population of killer whales. Of these, a substantial number (certainly >100) belong to the "transient" race (Barrett-Lennard et al. 1995), whose diet consists mainly of marine mammals, including Steller sea lions (Matkin et al. 2002, 2007). Unfortunately, very little is known about the spatial or temporal distribution of these whales, apart from a few isolated data points. What is known (from killer whale intake rates) is that as few as 27 male or 40 female killer whales switching to a 100% sea lion diet would have consumed enough animals to cause the observed decline (Williams et al. 2004). We therefore decided to use the abundance of the killer whales' more profitable alternative prey (harbor seals) as an inverse proxy for the Steller sea lions' predation mortality rate.

For the purposes of our model, we assumed that SSL pup survival (in H_9) and/or non-pup survival (in H_{10}) decrease(s) when harbor seal numbers near rookery i in year t, $h(i, t)$, fall below a critical level, h_{crit} (Fig. 3d):

$$\omega_n(i, t) = \begin{cases} 1 & \text{if } h(i, t) > h_{\text{crit}} \\ (1 - c_n) & \text{otherwise.} \end{cases} \quad (10)$$

If killer whale predation is a significant source of mortality for Steller sea lions, the average density of harbor seals around sea lion rookeries with rising populations should be higher than that around rookeries with falling populations. The existence of this relationship (Fig. 5) lends qualitative support to the hypothesis, though there may be alternative explanations yet to be

explored (Mangel and Wolf 2006). The critical harbor seal density is predicted to lie in between the two distributions. We therefore estimated h_{crit} as 4500, which is close to the midpoint between the harbor seal density averages for sea lion rookeries with rising and falling populations. Our work separates plausibility and evidential weight in support of this and all other hypotheses considered. For further details, see Mangel and Wolf (2006) and Wolf and Mangel (2007).

Our estimates of harbor seal density (Fig. 4d) came from online NMFS/AFSC marine mammal stock assessments and reports (Withrow et al. 2000, 2001, 2002, Angliss and Lodge 2002), a Marine Mammal Commission report (Hoover-Miller 1994), and eight journal articles (Bailey and Faust 1980, Everitt and Braham 1980, Pitcher 1990, Frost et al. 1999, Mathews and Pendleton 2000, Jemison and Kelly 2001, Boveng et al. 2003, Small et al. 2003). We constructed time series for nine regions and then estimated harbor seal abundance near each sea lion rookery as the sum of all nine region totals, each multiplied by the fraction of its seals estimated to be within 300 km of swimming distance from the rookery. We estimated the harbor seal non-pup total for a calendar year (using numbers from June and July, or occasionally August) and applied this estimate to the sea lion survival rates in the 12 months leading up to June 30 of that year.

Putting it all together

Each vital rate scaling function $\omega_n(i, t)$ (corresponding to hypothesis H_n) contains a parameter, c_n, that is unknown, but which we assumed is constant across all rookeries and years and which we intend to estimate from the data. The local value of each vital rate ($\rho_{i,t}$, $\sigma_{i,t}$, or $\phi_{i,t}$) was calculated by multiplying the maximum potential rate (ρ_0, σ_0, or ϕ_0) by all the relevant scaling functions. For each hypothesis, we wanted to know the MLE estimate of the corresponding c_n parameter and determine whether it was significantly different from zero, indicating that the data support the hypothesis.

Note that although the parameter values are fixed, the predicted strength of each factor at a particular rookery in a given year depends on the local conditions, as specified by the scaling functions. This allows different factors to dominate in different times and places, explaining different parts and segments of the decline.

The full model, which accounts for all 10 hypotheses, is one in which the life history parameters take the form:

$$\begin{aligned}\phi(i,t) &= \phi_0\omega_1(i,t)\omega_4(i,t)\\ \rho(i,t) &= \rho_0\omega_2(i,t)\omega_5(i,t)\omega_7(i,t)\omega_9(i,t)\\ \sigma(i,t) &= \sigma_0\omega_3(i,t)\omega_6(i,t)\omega_8(i,t)\omega_{10}(i,t).\end{aligned} \tag{11}$$

Calculating the likelihood

Starting with the beta-binomial observation error distribution (Appendix B) and a two-life-stage stochastic population model employing the local vital rates, we calculated the probability of observing the sequence of reported pup and non-pup counts at a particular rookery, given (1) relevant local conditions, and (2) a particular set of parameter values (c_1, c_2, …, c_{10}) in the hypothesized equations. We found this probability using backwards iteration and by taking the average across many possible true population trajectories that each could have produced the observed sequence of census data (Fig. 6). Details are discussed in the section *Averaging across population trajectories*.

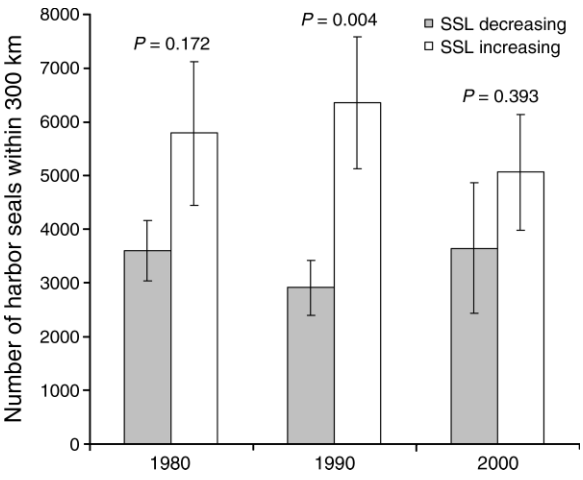


FIG. 5. Estimated harbor seal density (mean ± SE) around Steller sea lion rookeries with increasing or decreasing Steller sea lion (SSL) populations in 1980, 1990, and 2000. We used the apparent dividing line, ~4500 harbor seals, as a very rough estimate of the prey switching density.

Accounting for process uncertainty

Each of the life history parameters describes a transition probability between 0 and 1. The probability distributions for true numbers of non-pups and pups at time t, given the numbers at $t-1$, are computed from the associated binomial distributions. In this section, for simplicity, we drop subscripts and use $N(t)$ instead of $N_{true}(i, t)$. If we let $\Pr\{N(t)\,|\,N(t-1)\}$ denote the probability that the true non-pup population in year t is $N(t)$ given that the non-pup population in year $t-1$ was $N(t-1)$, then

$$\begin{aligned}&\Pr\{N(t)\,|\,N(t-1)\}\\ &= \sum_{s=0}^{\min[N(t-1),N(t)]} \left\{ \begin{bmatrix} N(t-1) \\ s \end{bmatrix} \sigma^s (1-\sigma)^{N(t-1)-s} \right.\\ &\qquad \times \begin{bmatrix} N(t-1) \\ N(t)-s \end{bmatrix} (\phi\rho)^{N(t)-s}\\ &\qquad \left. \times (1-\phi\rho)^{N(t-1)-[N(t)-s]} \right\}.\end{aligned} \tag{12}$$

Eq. 12 is a convolution of binomials that accounts for all the possible permutations of surviving non-pups [$s = 0, 1, 2, \ldots, N(t-1)$] and recruiting newborn pups [$N(t)$

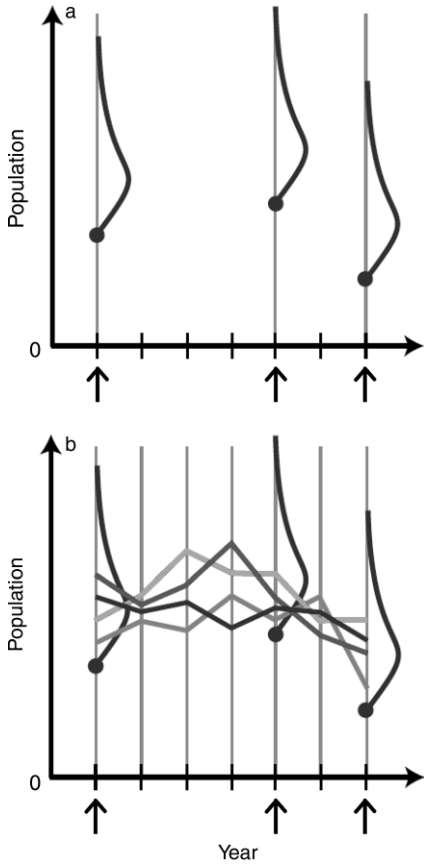


FIG. 6. The method of "thinking along sample paths," illustrated for the three census years marked by arrows. (a) Associated with each census (represented by the solid dot) is a probability distribution for the true number of non-pups, which is always larger than the observed number in the case of binomial or beta-binomial observation error. (b) Likelihood is calculated as the expectation, across all possible underlying trajectories (four are shown; gray lines), of the probability of observing the reported series of censuses.

$- s]$ that could have produced $N(t)$ non-pups at time t, given $N(t-1)$ non-pups at time $t-1$. The binomial transition probabilities are σ and $\phi\rho$, respectively.

If the number of pups at time $t-1$ is known, then we do not need to sum across possible pup numbers, and Eq. 12 becomes

$$\Pr\{N(t) = n \,|\, N(t-1), J(t-1)\} = \sum_{s=0}^{\min[N(t-1),N(t)]} \left\{ \begin{bmatrix} N(t-1) \\ s \end{bmatrix} \sigma^s (1-\sigma)^{N(t-1)-s} \times \begin{bmatrix} J(t-1) \\ N(t)-s \end{bmatrix} \rho^{N(t)-s} \times (1-\rho)^{J(t-1)-[N(t)-s]} \right\}. \quad (13)$$

Finally, the probability distribution for pups given the number of non-pups at time t is a single binomial:

$$\Pr\{J(t) \,|\, N(t)\} = \begin{bmatrix} N(t) \\ J(t) \end{bmatrix} \phi^{J(t)} (1-\phi)^{N(t)-J(t)}. \quad (14)$$

Averaging across population trajectories

We wanted to find the values of $c_1, c_2, \ldots, c_{10}$ that give the highest likelihood of observing all the reported SSL counts, using all of the data at hand. To calculate this likelihood, we took advantage of the rules of probability. In particular, if $\Pr\{A \,|\, B\}$ indicates the probability of the event A, given the event B, then the law of total probability gives

$$\Pr\{A \,|\, B\} = \sum_C \Pr\{A \,|\, C\} \Pr\{C \,|\, B\} \quad (15)$$

where the sum is taken over all the events (or states) C that can occur between B and A (Mangel and Clark 1988). In physics, this approach is known as path integration (Schulman 1981) and was first made very useful by Richard Feynman in his formulation of quantum mechanics (Feynman 1948). For that reason, we call it the method of "thinking along sample paths" (Mangel 2006). The method involves calculating the expectation of the likelihood across the range of possible underlying population trajectories implied by the observation error distribution and population model (Fig. 6).

Looking backward

The likelihood associated with a particular set of parameter values was estimated as the product of the likelihoods for all rookeries. The likelihood for an individual rookery was calculated using backwards iteration, which is an efficient way to consider many different possible "sample paths" simultaneously. First, we considered a range of possible values for $N_{\text{true}}(T)$, the true number of non-pups at the time of the latest available census, and calculated, conditioned on the parameters, the probability of observing the reported non-pup and pup counts ($N_{\text{obs}}(T)$ and $J(T)$) given $N_{\text{true}}(T)$ using Eqs. 2 and 14, respectively. Second, we moved to the previous year ($T-1$) and used Eqs. 12 or 13 to calculate, for a range of possible $N_{\text{true}}(T-1)$, the probability of making the transition to each $N_{\text{true}}(T)$ value considered in the first step. For each value of $N_{\text{true}}(T-1)$, the likelihood of observing $J(T)$ and $N_{\text{obs}}(T)$ is the sum of the probabilities calculated in the first step, weighted by the integrated binomial probabilities corresponding to each possible transition from $N_{\text{true}}(T-1)$ to $N_{\text{true}}(T)$. Formally expressed, we have

$$\Pr\{J(T), N_{\text{obs}}(T) \,|\, N_{\text{true}}(T-1)\} = \sum_{n=N_{\text{obs}}(T)}^{\infty} [\Pr\{N_{\text{obs}}(T) \,|\, N_{\text{true}}(T) = n\} \times \Pr\{N_{\text{true}}(T) = n \,|\, N_{\text{true}}(T-1)\} \times \Pr\{J(T) \,|\, N_{\text{true}}(T) = n\}] \quad (16)$$

where

$$\Pr\{N_{\text{obs}}(T) \,|\, N_{\text{true}}(T) = n\}$$

is calculated using Eq. 2,

$$\Pr\{N_{\text{true}}(T) = n \,|\, N_{\text{true}}(T-1)\}$$

is calculated using Eqs. 12 or 13, and

$$\Pr\{J(T) \,|\, N_{\text{true}}(T) = n\}$$

is calculated using Eq. 14.

We carried out the summation in Eq. 16 using steps of 20 over the range from $N_{\text{obs}}(T)$ to 8507, with this upper limit set to the highest estimated female non-pup count (6380) multiplied by a factor of $(\alpha+\beta)/\alpha$ (the inverse of the mean of the beta distribution). We chose a step size of 20 for N_{true} in order to minimize computer run-time without noticeably altering the shape of the output likelihood curves.

Analogous procedures allowed us to calculate the probabilities of observing all reported counts from $T-2$ onwards, given a range of possible values for $N_{\text{true}}(T-2)$. Then the process was repeated for $T-3$, $T-4$, etc., until we reached the earliest recorded census at t_{start}.

Model likelihood for a single rookery was calculated as the average of the likelihoods of all possible trajectories starting at each possible combination of observed true pup and unknown non-pup numbers, $J(t_{\text{start}})$ and $N_{\text{true}}(t_{\text{start}})$. The equal weight given to each trajectory likelihood in this average reflects our use of a flat (uniform) prior probability distribution for the unknown $N_{\text{true}}(t_{\text{start}})$. Overall model likelihood was calculated as the product of the likelihoods from all rookeries.

The likelihood of a model with parameters c_1, c_2,..., c_{10} given the census data is then:

$$L\{c_1, c_2, \ldots, c_{10} \,|\, \text{data}\} = \sum_{\substack{\text{all\_possible} \\ N_{\text{true}}(i,t_{\text{start}})}} \left[\Pr\{N_{\text{true}}(i,t_{\text{start}})\} \prod_{\text{all censuses}} \times \Pr\{N_{\text{obs}}(i,t), J(i,t) \,|\, N_{\text{true}}(i,t_{\text{start}}), c_1, c_2, \ldots, c_{10}, \alpha, \beta, \text{etc.}\} \right]. \quad (17)$$

We searched the 10-dimensional parameter space to find the joint maximum likelihood estimate (MLE) of the parameters and then calculated the marginal likelihood curve for each parameter by taking the average across all other dimensions for which the MLE estimate was nonzero. According to Bayes' theorem, the resultant marginal likelihood may be normalized and interpreted as a posterior probability if we assume a uniform prior for each parameter. This allowed us to calculate a confidence interval around the MLE value, which we used to determine whether each parameter is significantly different from zero.

We will say that the data provide *strong support* for a hypothesis if the MLE of the parameter associated with that hypothesis is nonzero and its confidence interval does not include zero. If the MLE is 0, the data provide *no support* for the hypothesis. Furthermore, we classified the supported hypotheses as having a strong effect or a moderate effect, depending upon how the MLE value of the parameter affects the predicted population dynamics of Steller sea lions.

Missing production: an alternative way to look at the decline and the hypotheses

Using the background vital rates (Eq. 4) and a deterministic version of the two-stage population model, we can calculate the expected number of animals in year t, given the number that were counted in the previous year. The difference between this and the number actually counted in year t is the number of missing non-pups, $M_0(i, t)$, which is what we hoped to explain:

$$M_0(i,t) = N_{\text{obs}}(i,t-1)(\phi_0\rho_0 + \sigma_0) - N_{\text{obs}}(i,t). \quad (18)$$

We expanded this method to non-census years by using log-linear interpolation to estimate N_{obs} between censuses. Then we could plot the estimate of missing animals for each rookery and every year, producing a "landscape" of missing animals. After we had MLE values for all 10 c_n parameters, we could estimate the number of animals lost due to all hypothesized effects:

$$M_{1,2,3,\ldots,10}(i,t) = N_{\text{obs}}(i,t-1)(\phi_0\rho_0 + \sigma_0) - N_{\text{obs}}(i,t-1)[\phi(i,t-1)\rho(i,t) + \sigma(i,t)] \quad (19)$$

where $\phi(i, t)$, $\rho(i, t)$, and $\sigma(i, t)$ are defined in Eq. 11.

The lost production due to the effects of individual hypotheses may be calculated by setting all other ω_n values (in Eq. 11) to 1, which is equivalent to setting the respective c_n values to zero.

Differences from a previous version of the model

This work is an extension of an earlier model (Wolf 2006, Wolf et al. 2006) that used more general functional forms for H_1–H_3 and H_7–H_8 and a less precise dataset for fishery activity, along with many other minor differences. The previous analysis also differed in using profile likelihoods rather than full marginal likelihoods.

Results

We found strong support for four hypotheses: H_2, H_4, H_5, and H_{10}. The MLE values of the corresponding parameters (c_2, c_4, c_5, and c_{10}) were all nonzero and none of their 95% confidence intervals included zero (Table 2). The marginal log-likelihood plots for these parameters are shown in Fig. 7. The data do not support any of the

TABLE 2. Maximum likelihood estimates (MLE) and confidence intervals (based on area under the curve) for the parameters associated with the four supported hypotheses.

Hypothesis	MLE	95% confidence interval
H_2: lower prey density → less pup recruitment	2 380 000	2 276 000–2 484 000
H_4: higher pollock fraction → lower fecundity rate	0.053	0.0236–0.0824
H_5: higher pollock fraction → less pup recruitment	3.67	3.312–4.028
H_{10}: fewer HS (more predation) → lower non-pup survival	0.023	0.01816–0.02784

other six hypotheses (MLE values for c_1, c_3, c_6, c_7, c_8, and c_9 were all zero).

In Fig. 8a–d we show the vital rate scaling functions for H_2, H_4, H_5, and H_{10} using the corresponding MLE values of c_2, c_4, c_5, and c_{10}, respectively. H_2 (Fig. 8a) predicts a steep drop in pup recruitment when total prey density drops below about 2 000 000 kg/km^3. H_4 (Fig. 8b) predicts a gradual decline in fecundity as the fraction of pollock in available prey increases, topping out at a 10% reduction when 85% of the available prey biomass is pollock. H_5 (Fig. 8c) predicts a massive decline in pup recruitment as the pollock fraction increases, with about a 90% reduction when 50% of the available prey biomass is pollock. H_{10} (Fig. 8d) predicts a 2–3% drop in non-pup survival whenever local harbor seal density falls below the critical level.

In Table 3, we give the Akaike Information Criterion (AIC) and AIC weight for five of the most likely model

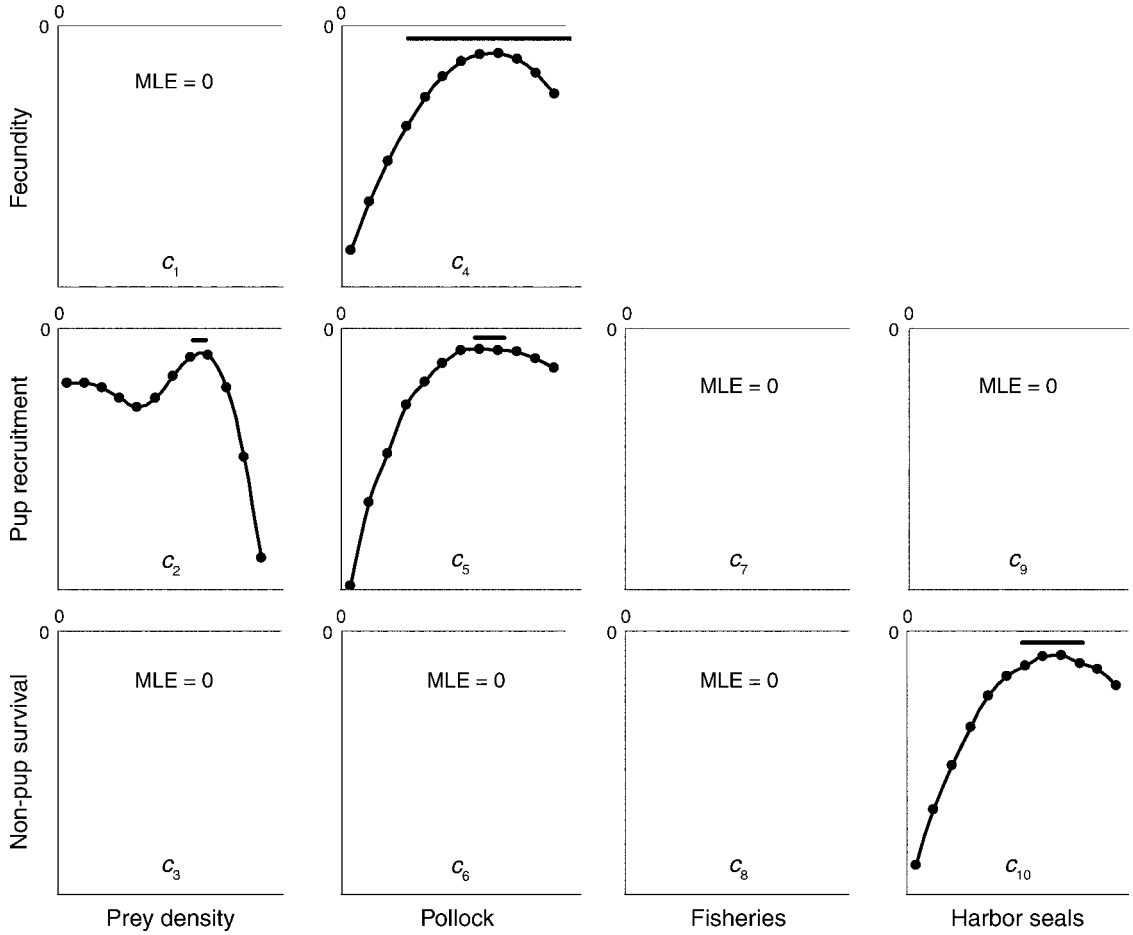


FIG. 7. Marginal log-likelihood curves. Each column corresponds to an environmental factor (prey density, pollock, fisheries, harbor seals), and each row indicates a vital rate affected (fecundity, pup recruitment, non-pup survival). The MLE (maximum likelihood estimate) values of c_2, c_4, c_5, and c_{10} are significantly different from zero, indicating strong support for H_2, H_4, H_5, and H_{10}, respectively. The solid bars above each curve indicate 95% confidence intervals (zones containing 95% of the total posterior probability). The curves for c_1, c_3, c_6, c_7, c_8, and c_9 (not shown) all have their likelihood peaks at zero, indicating that none of the corresponding hypotheses is supported by the data.

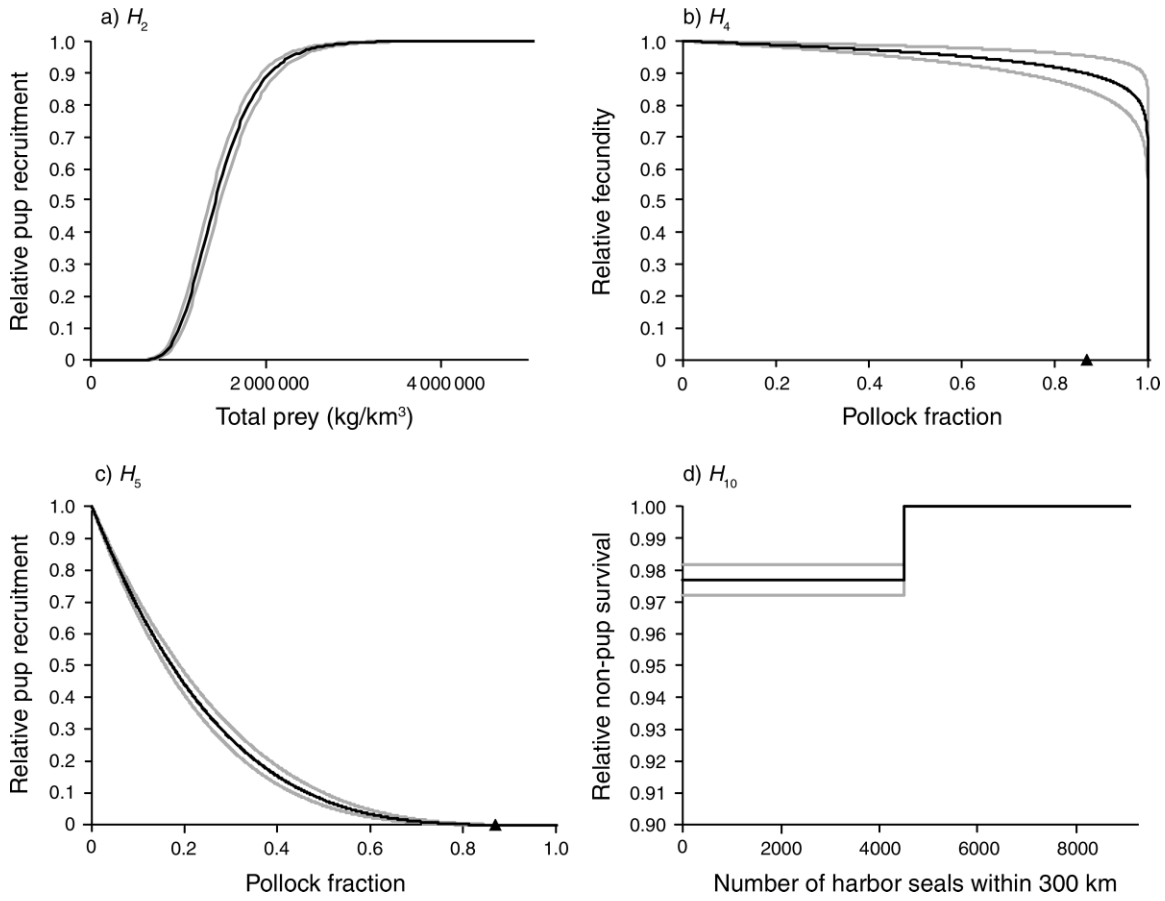


FIG. 8. Vital rate scaling functions corresponding to the MLE values of c_2, c_4, c_5, and c_{10}, along with their 95% confidence intervals. (a) Pup recruitment scaling function corresponding to H_2 and c_2. (b) Fecundity scaling function corresponding to H_4 and c_4. (c) Pup recruitment scaling function corresponding to H_5 and c_5. (d) Non-pup survival scaling function corresponding to H_{10} and c_{10}.

configurations. The model including H_2, H_4, H_5, and H_{10} carries 98% of the total weight, and essentially all of the remaining weight is carried by the model including H_2, H_5, and H_{10}. The models with only two hypotheses considered (regardless of which hypotheses) have essentially an AIC weight of zero.

Missing production

In Fig. 9a, we show the "landscape" of missing animals according to the background vital rates (Eq. 18). This is the pattern we wish to explain. Fig. 9b shows the expected landscape of missing animals according to the combined effects of H_2, H_4, H_5, and H_{10}, calculated using the MLE values of c_2, c_4, c_5, and c_{10}. Note that the obvious qualitative visual correspondence between these figures is an emergent property of the underlying model fit, not the reverse (in contrast with more conventional regression-based approaches).

We can also separate the different factors, calculating the impact of each supported effect in isolation. The equations for lost production due to H_2, H_4, H_5, and H_{10} (Fig. 10a–d, respectively) are as follows:

TABLE 3. Model likelihoods and AIC weights.

Model	No. parameters	Log likelihood	AIC	ΔAIC	AIC weight
c_2, c_4, c_5, c_{10}	4	−6833.593	13 675.186	0	0.98035323
c_2, c_5, c_{10}	3	−6838.503	13 683.006	7.82	0.01964677
c_4, c_5, c_{10}	3	−6873.966	13 753.932	78.746	7.79661×10^{-18}
c_2, c_4, c_5	3	−6901.473	13 808.946	133.76	8.82608×10^{-30}
c_2, c_4, c_{10}	3	−7225.786	14 457.572	782.386	1.2544×10^{-170}

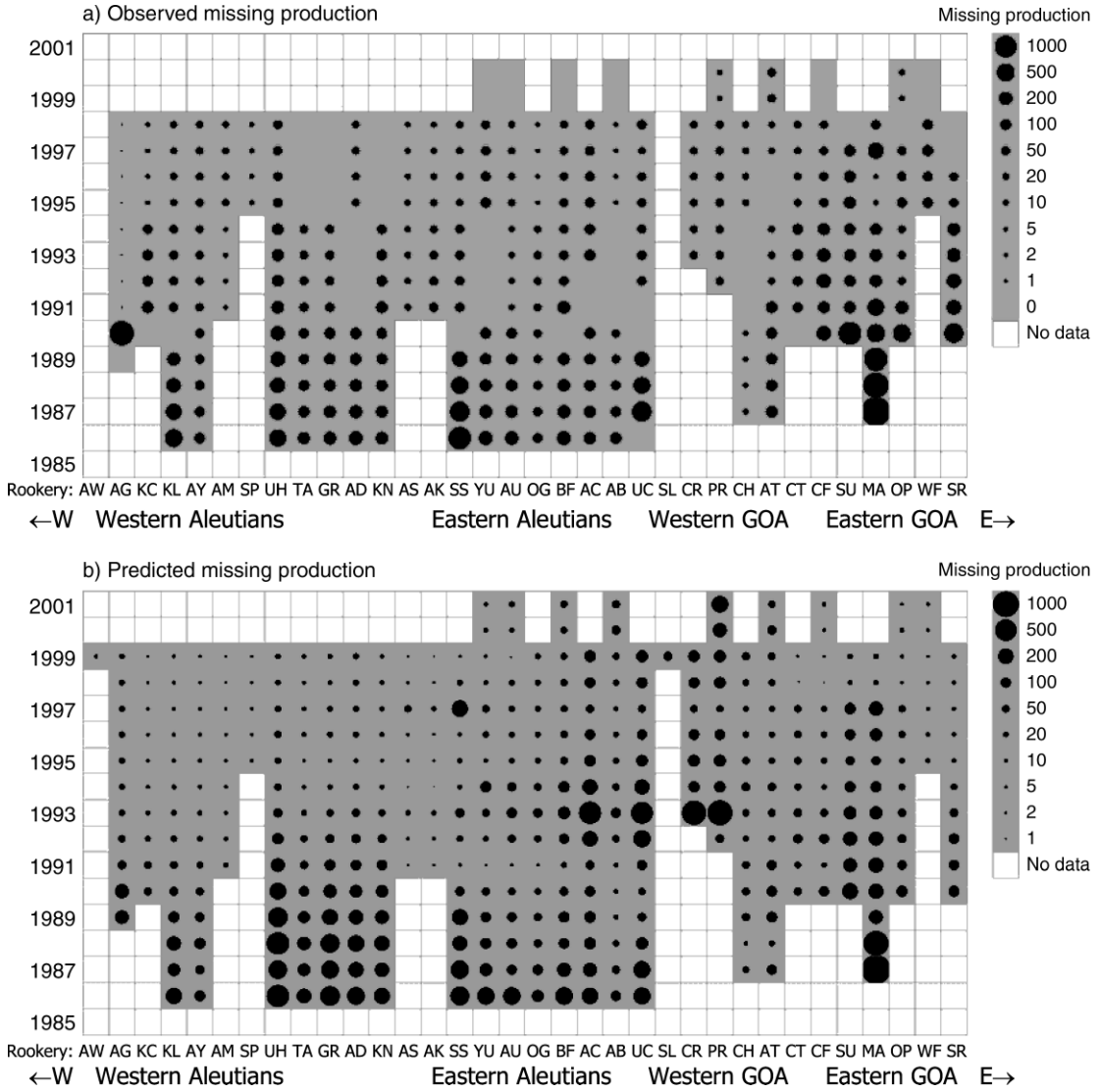


FIG. 9. Space–time plots of (a) observed and (b) predicted missing production (relative to pre-decline vital rates), according to Eq. 18. Missing production is defined as the number of animals that must be subtracted from a rookery in a particular year in order for its modeled population trajectory to match the observed counts. Panel (a) shows the pattern we set out to explain. Panel (b) shows the combined effect of the four supported hypotheses H_2, H_4, H_5, and H_{10}, calculated using Eq. 19 and the respective MLE values for c_2, c_4, c_5, and c_{10}. See Appendix A: Table A1 for full rookery names.

$$M_2(i,t) = N_{\text{obs}}(i,t-1)(\phi_0\rho_0 + \sigma_0) - N_{\text{obs}}(i,t-1)[\phi_0\rho_0\omega_2(i,t) + \sigma_0] \quad (20)$$

$$M_4(i,t) = N_{\text{obs}}(i,t-1)(\phi_0\rho_0 + \sigma_0) - N_{\text{obs}}(i,t-1)[\phi_0\omega_4(i,t-1)\rho_0 + \sigma_0] \quad (21)$$

$$M_5(i,t) = N_{\text{obs}}(i,t-1)(\phi_0\rho_0 + \sigma_0) - N_{\text{obs}}(i,t-1)[\phi_0\omega_5(i,t-1) + \sigma_0] \quad (22)$$

$$M_{10}(i,t) = N_{\text{obs}}(i,t-1)(\phi_0\rho_0 + \sigma_0) - N_{\text{obs}}(i,t-1)[\phi_0\rho_0 + \sigma_0\omega_{10}(i,t)]. \quad (23)$$

The formula for the combined effect of all four (Fig. 9b) is

$$M_{2,4,5,10}(i,t) = N_{\text{obs}}(i,t-1)(\phi_0\rho_0 + \sigma_0) - N_{\text{obs}}(i,t-1) \times [\phi_0\omega_4(i,t-1)\rho_0\omega_2(i,t)\omega_5(i,t) + \sigma_0\omega_{10}(i,t)]. \quad (24)$$

The decline as an emergent property

Though our model was fitted at the level of individual rookeries and years, it should produce an aggregate, range-wide pattern that matches the observed historical decline. We tested this by running 1000 forward

simulations (1985 to 1999) of the 28 rookeries in the western population for which the 1985 population was known or could be estimated by log-linear interpolation. The range of the 1000 composite trajectories is shown in Fig. 11, and compared with the observed data from the same period, summed across the same 28 rookeries. Like the comparison between Fig. 9a, b above, the visual correspondence in Fig. 10 confirms that the model's emergent pattern qualitatively matches the data.

Discussion

The strong message of ecology is that change is ubiquitous and that the reasons for change are manifold. Thus, rather than trying to "prove" one mechanism, we should recognize that multiple mechanisms will almost always be at work, and we should ask how to weigh the importance of different mechanisms. It is this approach that we have taken in understanding the decline of the western population of Steller sea lions. Based on the best set of data examined to date, we conclude that there is good evidence for two strong effects (H_2 and H_5) and two moderate effects (H_4 and H_{10}), and no evidence for any of the other six effects considered (Tables 1 and 2).

Virtually all of the AIC weight (98%) is assigned to the model with H_2, H_4, H_5, and H_{10}, suggesting that these four hypotheses (or analogous ones that have the same apparent behavior) are sufficient to explain the decline. There is always the possibility that, if we had included additional hypotheses beyond the 10 tested in this paper, we might have found others with significant support from these data. Nevertheless, we believe that we have tested the hypotheses that represent all of the current "best guesses" with sufficient rookery-scale data to permit their evaluation.

As shown in Fig. 10, each of the four supported hypotheses corresponds to a different pattern of impact. H_2 and H_5 (the effects of prey density and pollock fraction on pup recruitment) have by far the strongest impact, but the other two are also important at certain rookeries and in certain years. There is some spatial and temporal overlap between the impacts of the different factors, but as the figures show and the AIC analysis confirms, all four are necessary to approximate the observed pattern of the decline in Fig. 9a.

Our predictions compare well with those reported by Guenette et al. (2006) in their simulation-based study of a section of the declining SSL population. In their study, the two strongest factors were found to be shifting prey species composition (due to the Pacific decadal oscillation) and predation by killer whales. Our model also predicts strong negative impacts of prey species composition (H_5; pollock fraction) and killer whale predation (H_{10}) in the same area (central and western Aleutians), as shown by Figs. 10b and 9d, respectively. The Guenette et al. (2006) study also found weaker effects of fisheries activity and competition between SSL and halibut, both factors involving the depletion of Atka mackerel. This depletion is captured in our model as prey abundance, for which we found a strong effect (H_2; Fig. 10a).

The predicted negative effects of prey abundance and prey species composition on population trends, particularly in the early part of the SSL decline (H_2, H_4, and H_5; Fig. 10a–c), agree well with previous studies that found the 1980s to be characterized by elevated mortality rates for young sea lions (York 1994, Holmes and York 2003), along with low body mass (Calkins et al. 1998) and other signs of nutritional stress (Trites and Donnelly 2003). The strong effect of pollock fraction (H_4, H_5; Fig. 10b, c) also confirms the finding of Merrick et al. (1997) that diet diversity correlates with population growth rate. However, our predictions regarding fecundity do not match those of Holmes and York (2003) and Holmes et al. (2007) particularly well. They found evidence for depressed fecundity rates in the 1990s in the central Gulf of Alaska, whereas our model predictions show only a small amount of lost fecundity in this period, somewhat further to the west (H_4; Fig. 10b). The reason for this difference is not known.

From pattern to mechanism

The pattern of lost production across time and space (Fig. 9a) reveals a startling degree of spatial and temporal heterogeneity in the decline that was completely obscured in the composite dataset (Fig. 1). The "landscapes" of missing animals associated with each effect (Fig. 10) drive home the point that different factors can be important in different places and at different times. Rather than a single "smoking gun," we have four. But what are the actual mechanisms underlying each effect?

H_2 (prey abundance affects pup recruitment) suggests a simple mechanism: Low abundance of prey (excluding pollock) leads to diminished pup recruitment rates, presumably because pups are less adept at catching prey than older sea lions and more likely to experience a foraging shortfall.

The mechanisms behind H_4 (pollock fraction affects fecundity) and H_5 (pollock fraction affects pup recruitment) are more ambiguous. A high fraction of pollock in the environment can result either from high pollock abundance or from low non-pollock abundance, implying two alternative mechanisms. The plausibility of each mechanism depends in part on the degree to which Steller sea lions are generalist predators, consuming prey in proportion to availability. In the first mechanism, overabundance of pollock leads to overconsumption of this low-quality food, causing malnutrition, which impacts fecundity or pup survival probabilities. In the second mechanism, pups or pregnant females avoid pollock and require a particular prey type that scales inversely with it, risking starvation when the alternative prey is in short supply. However, as mentioned previously, we would expect this effect to appear primarily as support for H_1 or H_2. Therefore, we conclude that the observed support for H_4 and H_5 is

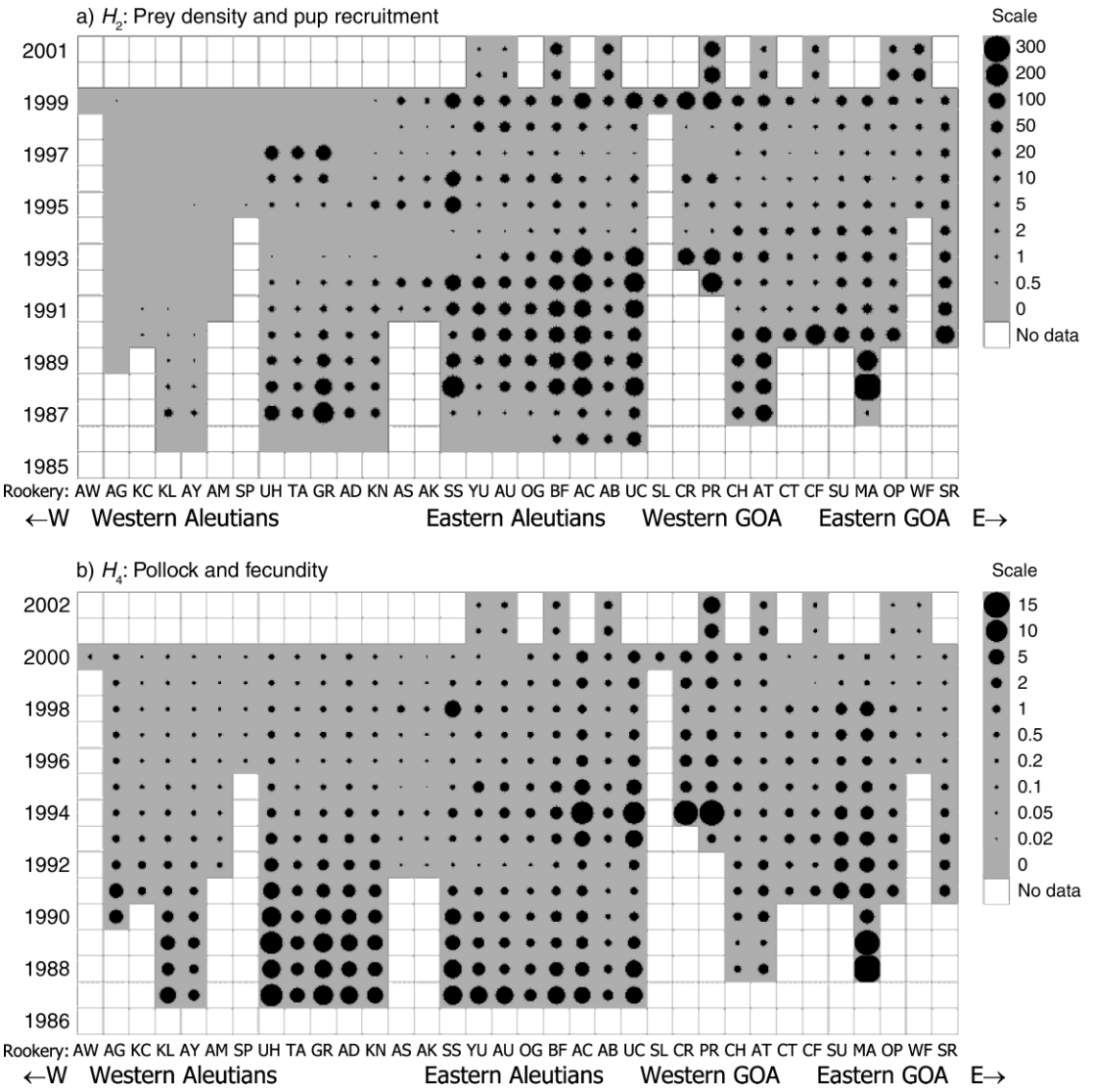


FIG. 10. Space–time plots of missing production based on Eqs. 20–23 for (a) H_2 (prey density affecting pup recruitment), (b) H_4 (pollock fraction affecting fecundity), (c) H_5 (pollock fraction affecting pup recruitment), and (d) H_{10} (harbor seal density affecting non-pup survival), respectively. See Appendix A: Table A1 for full rookery names.

more indicative of a problem with the prey field being dominated by pollock (as in Rosen and Trites 2000), rather than a problem with too little non-pollock. It is not known why the sea lions would fail to exclude such an apparently suboptimal prey type from their diets.

Our finding of strong support for both H_2 and H_5 suggests a possible synergy between the effects of low prey availability and high pollock fraction on pup survival: In light of the limited digestive capacity of pups and the low energy density of pollock (Rosen and Trites 2000, 2004), it seems quite plausible that a series of unsuccessful foraging bouts broken only by encounters with pollock would prove fatal. Even if they ate pollock until they were full, the young sea lions might not be able to "recharge" themselves under these conditions. This idea agrees well with that proposed by Rosen and Trites (2000, 2004), based upon their studies of captive Steller sea lions that were fasted for short periods and/or allowed to consume only low-energy prey.

H_{10} had a small but significant effect, reducing annual survival of non-pups by ~2% only at sea lion rookeries where the local harbor seal population is below 4500 individuals. We have suggested a mechanism involving optimal prey switching by killer whales (Eq. 10; see also Mangel and Wolf 2006), but clearly there may be alternative explanations for the association of low harbor seal numbers with negative population growth in Steller sea lions (Fig. 5). For example, both species may be affected by a common local factor such as disease or predation. The topic has already generated

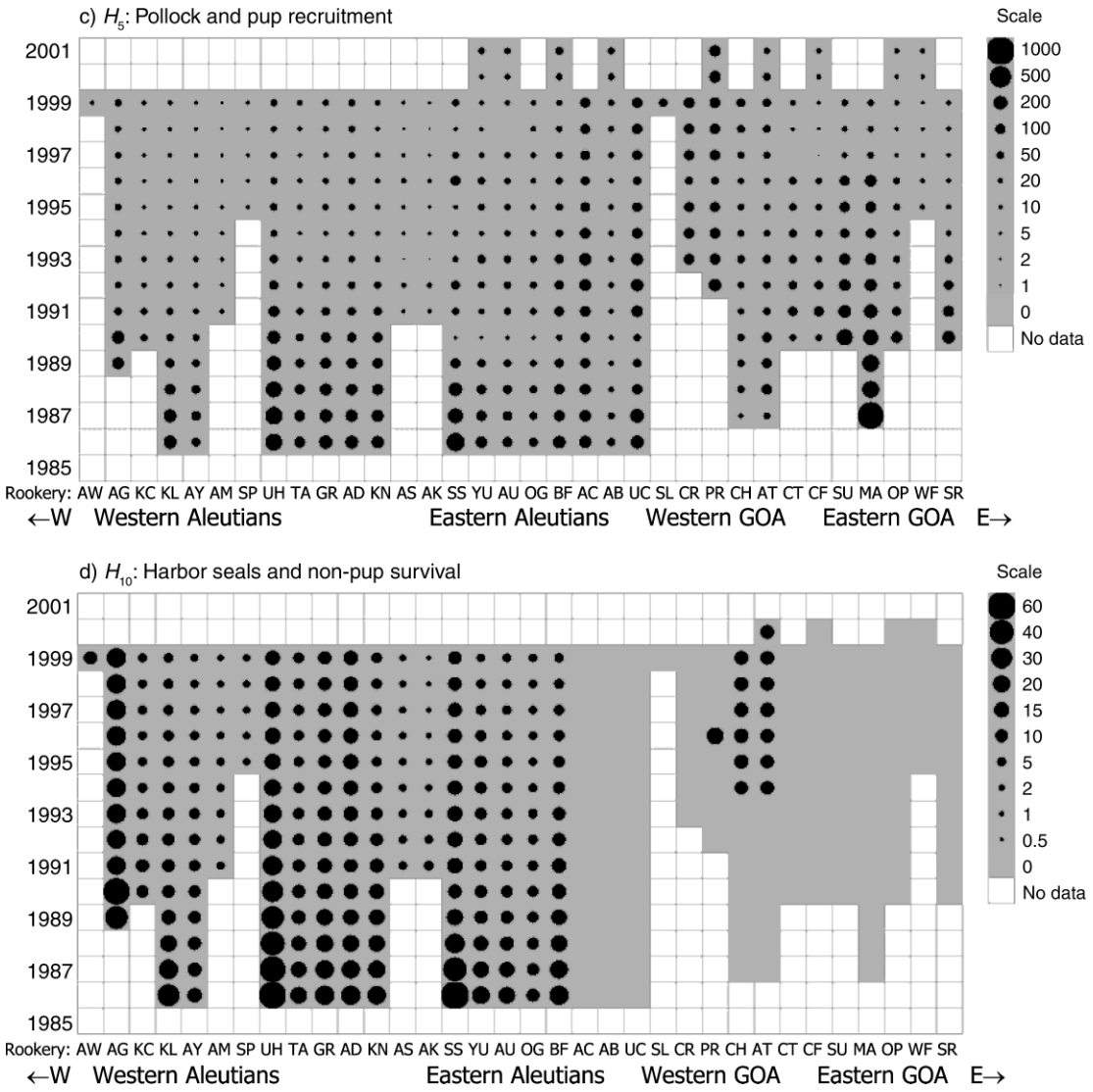


FIG. 10. Continued.

considerable controversy (for a recent review, see DeMaster et al. 2006). Whatever the mechanism may be, it is clearly worthy of further investigation because it explains some residual variation in Steller sea lion vital rates that is not explained by prey availability, pollock fraction, or fisheries impacts.

Adaptive management

The results of a model such as ours provide valuable insight, but they also reveal new questions and direct further research. In the current investigation, our results suggest an adaptive management plan in which one designates the areas around some of the rookeries as experimental zones where the fishery quotas are contingent upon the results of pre-fishing-season survey trawls. One might envision a series of treatments: (1) rookeries around which fishing occurs (control type 1), (2) rookeries around which no fishing occurs (control type 2), (3) rookeries around which fishing is reduced or prohibited if the total prey biomass in the pre-season zone is below a critical threshold (determined by c_2), and (4) rookeries around which a directed pollock fishery occurs if the pre-season survey suggests pollock fraction is above a critical threshold (determined by c_5).

This combination of rookery types would allow sufficient variation in treatment, which is crucial in adaptive management (Parma et al. 1998). Sea lion vital rates would be monitored in the same areas to determine whether the management plan is having a positive effect. Our model could be adapted to simulate forward and suggest a time scale over which results might be expected to appear.

Our results also suggest a form of "adaptive observation": identify rookeries with high and low

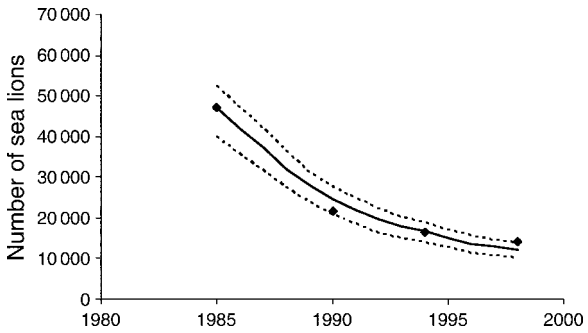


FIG. 11. Distribution of range-wide population trajectories for western Alaska, created using forward simulation and summed across the 28 rookeries for which 1985 counts were available or could be estimated by log-linear interpolation. The dotted lines are 1% and 99% quantiles of the 1000 simulation runs. The four diamonds show the observed counts, summed across the same 28 rookeries.

numbers of harbor seals (regardless of the number of sea lions). The prediction of H_{10} is that the per-capita attack rate of killer whales on sea lions will be higher around rookeries where harbor seal densities are low. Careful monitoring of killer whale attack rates would provide a natural test of Hypothesis 10. If the prediction is not confirmed (that is, if low harbor seal numbers are associated with declining sea lions, but not with elevated killer whale attack rates), then there must be some other factor to explain the observed correspondence between low harbor seals and declining sea lions.

Ecological detection

The purpose of ecological detection is to confront each putative mechanism of the decline with the data and allow the data to arbitrate between the different models on a level playing field. Ecological detection recognizes that our understanding of the world will always be incomplete, and that the goal should be to achieve the best understanding possible given the available data. The paradigm emphasizes the need to understand the role of multiple mechanisms in the decline of the Steller sea lion, rather than trying to explain the entire phenomenon with any single hypothesis.

There are two reasons why it is so important to consider multiple hypotheses simultaneously. First, it is very common for more than one hypothesis to be important in the system, and the nature of any interactions between different effects will not be captured if they are considered individually. Second, one can often find spurious support for any single hypothesis when it is considered in isolation. The only solution is to jointly solve for the strengths of all hypotheses at the same time, so that the dominant ones may emerge and leave little for the spurious ones to explain. Underlying this, and key to success, is finding the means to express all of the hypotheses within a common statistical currency (the parameters $\{c_n\}$).

In the case of the Steller sea lion model, when any of the 10 hypotheses are considered in isolation (with the other nine effects set to zero), the maximum likelihood estimate (MLE) of the relevant parameter is nonzero. In other words, all 10 of the hypotheses appear to fit the data when they are tested individually. But when they are all tested simultaneously, with the model searching for the joint maximum likelihood solution of all 10 parameters, all but the four listed above have their MLEs at zero, indicating that they have no effect. This observation serves as a reminder of the fact that likelihood is relative, and the apparent fit of any single hypothesis to the data does not rule out the possibility that other hypotheses may exist that have an even better fit. For example, we were not able to test hypotheses related to diseases, other pathogens, or negative interactions between sea lions and small boats (those lacking observers). If the fine-scale data necessary to evaluate these became available in the future, our model could be expanded to include the additional effects.

An illustration of the relativity of likelihood is provided by Hennen's (2006) study, in which he compared the population growth rates at individual Steller sea lion rookeries with the amount of commercial fishing activity that occurred within 20 km of each rookery. Using the same data that he later provided to us for the present study, Hennen found strong support for a negative effect of fishing activities on sea lions before 1991. But in the context of the full sea lion model described here, when the hypotheses describing the effects of fishing activities on pup recruitment and non-pup survival are allowed to "compete" against the other eight hypotheses, their MLE values are zero, indicating no support for those effects. There is some possibility that we might have found statistical support for this factor if the fisheries data had not been pooled into two broad time steps, but the complete disappearance of support between the previous single-hypothesis study and ours suggests that the other factors simply fit the observed pattern more closely. It is also possible that a different spatial scale or an entirely different index would have better captured the impact of humans on Steller sea lions (perhaps due to illegal shooting; see NRC 2003), but such speculation is beyond the scope of the current study.

CONCLUSIONS

During the decades-long search for the cause or causes of the Steller sea lions' decline, the question "Is it food?" has been asked a number of times (Alaska Sea Grant College Program 1993). As with most questions in biology, we shall never be able to *prove* that food is a major factor. However, we conclude that the weight of the current evidence is that food did indeed cause the majority of the problem: And both the quantity and quality of the food matters. The more recent question "Is it killer whale predation?" can be answered too: Yes, the data are consistent with a significant impact of killer

whales on Steller sea lions at times and places where the harbor seal density is low. However, additional research is needed in order to confirm that this is the true mechanism.

These results are not entirely unexpected (indeed, each of the 10 hypotheses is plausible and has been proposed at some point, with associated supporting data). What our work has done is to use the weight of the evidence, when all plausible hypotheses are competing, to find those that jointly win the competition.

Acknowledgments

We thank Lowell Fritz, Anne York, Tim Ragen, Doug DeMaster, Gunnar Steffanson, Jim Estes, Andrew Trites, Graeme Parkes, Andrew Constable, the Parker and Mangel labs, Akeena Solar, and Katrina Dlugosch for extremely useful comments and support at various stages of the project. The manuscript benefited greatly from comments made by two anonymous reviewers. We also thank Dan Hennen for graciously providing his compiled fishery activity data. N. Wolf heartily thanks Hanna Kokko, Ian Boyd, Jarl Giske, Rune Rosland, Christian Jorgensen, Oyvind Fiksen, Sigrunn Eliasson, and other members of the modeling group in the University of Bergen's Biology department for their academic encouragement. Funding in the early stages was provided by the Alaska Fisheries Science Center via contract number AB133F-02-CN-0085, administered through MRAG Americas. The work was also partially supported by NSF grant DMS 03-10542 to Marc Mangel and by the Center for Stock Assessment Research, a partnership between the Santa Cruz Laboratory of the National Marine Fisheries Service and UCSC.

Literature Cited

Abramowitz, M., and I. A. Stegun. 1964. Handbook of mathematical functions with formulas, graphs, and mathematical tables. Dover, New York, New York, USA.

Alaska Sea Grant College Program. 1993. Is it food? Addressing marine mammal and seabird declines: workshop summary. AK-SG-93-01. University of Alaska, Alaska Sea Grant College Program, Fairbanks, Alaska, USA.

Alverson, D. L. 1992. A review of commercial fisheries and the Steller sea lion (*Eumetopias jubatus*): the conflict arena. Reviews in Aquatic Sciences 6:203–256.

Anderson, P. J., and J. E. Blackburn. 2002. Status of demersal and epibenthic species in the Kodiak Island and Gulf of Alaska region. Pages 57–60 *in* D. DeMaster and S. Atkinson, editors. Steller sea lion decline: Is it food II? University of Alaska Sea Grant College Program, Fairbanks, Alaska, USA.

Andrews, R. D., D. G. Calkins, R. W. Davis, B. L. Norcross, K. Peijnenberg, and A. W. Trites. 2002. Foraging behavior and energetics of adult female Steller sea lions. Pages 19–22 *in* D. DeMaster and S. Atkinson, editors. Steller sea lion decline: Is it food II? University of Alaska Sea Grant College Program. Fairbanks, Alaska, USA.

Angliss, R. P., and K. L. Lodge. 2002. Alaska marine mammal stock assessments, 2002. NOAA Technical Memorandum, National Marine Fisheries Service AFSC-133. U.S. Department of Commerce, National Oceanic and Atmospheric Administration, National Marine Fisheries Service, Alaska Fisheries Science Center, Seattle, Washington, USA.

Bailey, E. P., and N. H. Faust. 1980. Summer distribution and abundance of marine birds and mammals in the Sandman Reefs, Alaska, USA. Murrelet 61:6–19.

Barrett-Lennard, L. G., K. A. Heise, E. Saulitis, G. Ellis, and C. Matkin. 1995. The impact of killer whale predation on Steller sea lion populations in British Columbia and Alaska. North Pacific Universities Marine Mammal Research Consortium. Vancouver, British Columbia, Canada.

Bickham, J. W., T. R. Loughlin, D. G. Calkins, J. K. Wickliffe, and J. C. Patton. 1998. Genetic variability and population decline in Steller sea lions from the Gulf of Alaska. Journal of Mammalogy 79:1390–1395.

Boveng, P. L., J. L. Bengtson, D. E. Withrow, J. C. Cesarone, M. A. Simpkins, K. J. Frost, and J. J. Burns. 2003. The abundance of harbor seals in the Gulf of Alaska. Marine Mammal Science 19:111–127.

Burnham, K. P., and D. R. Anderson. 2002. Model selection and multimodel inference: a practical information-theoretic approach. Second edition. Springer, New York, New York, USA.

Calkins, D. G., E. F. Becker, and K. W. Pitcher. 1998. Reduced body size of female Steller sea lions from a declining population in the Gulf of Alaska. Marine Mammal Science 14:232–244.

Calkins, D., and K. W. Pitcher. 1982. Population assessment, ecology and trophic relationships of Steller sea lions in the Gulf of Alaska. Pages 447–546 *in* Environmental assessment of the Alaskan continental shelf, final report. U.S. Department of Commerce and U.S. Department of the Interior, Washington, D.C., USA.

Castellini, M. 1993. Report of the Marine Mammal Working Group. Pages 4–11 *in* S. Keller, editor. Is it food? Addressing marine mammal and seabird declines. Alaska Sea Grant College Program, Fairbanks, Alaska, USA.

Caughley, G. 1994. Directions in conservation biology. Journal of Animal Ecology 63:215–244.

Clark, C. W., and M. Mangel. 2000. Dynamic state models in ecology. Methods and applications. Oxford University Press, New York, New York, USA.

DeMaster, D. P., A. W. Trites, P. Clapham, S. Mizroch, P. Wade, R. J. Small, and J. V. Hoef. 2006. The sequential megafaunal collapse hypothesis: Testing with existing data. Progress in Oceanography 68(2–4):329–342.

Evans, M., N. Hastings, and J. B. Peacock. 2000. Statistical distributions. John Wiley. New York, New York, USA.

Everitt, R. D., and H. W. Braham. 1980. Aerial survey of pacific harbor seals, *Phoca vitulina richardsi*, in the southeastern Bering Sea. Northwest Science 54:281–288.

Ferrero, R. C., and L. W. Fritz. 2002. Steller Sea lion research and coordination: a brief history and summary of recent progress. NOAA Technical Memorandum NMFS-AFSC-129. U.S. Department of Commerce, National Oceanic and Atmospheric Administration, National Marine Fisheries Service, Alaska Fisheries Science Center, Anchorage, Alaska, USA.

Feynman, R. P. 1948. Space-time approach to non-relativistic quantum mechanics. Reviews of Modern Physics 20:367–387.

Frost, K. J., L. F. Lowry, and J. M. Ver Hoef. 1999. Monitoring the trend of harbor seals in Prince William Sound, Alaska, after the Exxon Valdez oil spill. Marine Mammal Science 15:494–506.

Gerber, L. R., and G. R. Van Blaricom. 2001. Implications of three viability models for the conservation status of the western population of Steller sea lions (*Eumetopias jubatus*). Biological Conservation 102:261–269.

Green, R. E. 1995. Diagnosing causes of bird population declines. Ibis 137:S47–S55.

Guenette, S., S. J. J. Heymans, V. Christensen, and A. W. Trites. 2006. Ecosystem models show combined effects of fishing, predation, competition, and ocean productivity on Steller sea lions (*Eumetopias jubatus*) in Alaska. Canadian Journal of Fisheries and Aquatic Sciences 63:2495–2517.

Hansen, D. J. 1997. Shrimp fishery and capelin decline may influence decline of harbor seal (*Phoca vitulina*) and northern sea lion (*Eumatopias jubatus*) in western Gulf of Alaska. Pages 197–207 *in* Proceedings of the International Symposium on the Role of Forage Fishes in Marine Ecosystems.

University of Alaska Sea Grant College Program Report Number 97-01. University of Alaska, Fairbanks, Alaska, USA.

Hennen, D. 2006. Associations between the Alaska Steller sea lion decline and commercial fisheries. Ecological Applications 16:704–717.

Hilborn, R., and M. Mangel. 1997. The ecological detective: confronting models with data. Princeton University Press, Princeton, New Jersey, USA.

Hoffman, J. I., C. W. Matson, W. Amos, T. R. Loughlin, and J. W. Bickham. 2006. Deep genetic subdivision within a continuously distributed and highly vagile marine mammal, the Steller's sea lion (*Eumetopias jubatus*). Molecular Ecology 15:2821–2832.

Holling, C. S. 1959. Some characteristics of simple types of predation and parasitism. Canadian Entomologist 91:385–398.

Holmes, E. E., L. W. Fritz, A. E. York, and K. Sweeney. 2007. Age-structured modeling reveals long-term declines in the natality of western Steller sea lions. Ecological Applications 17:2214–2232.

Holmes, E. E., and A. E. York. 2003. Using age structure to detect impacts on threatened populations: A case study with Steller sea lions. Conservation Biology 17:1794–1806.

Hoover-Miller, A. A. 1994. Harbor seal (*Phoca vitulina*) biology and management in Alaska. Contract Number T75134749. Marine Mammal Commission, Washington, D.C., USA.

Jemison, L. A., and B. P. Kelly. 2001. Pupping phenology and demography of harbor seals (*Phoca vitulina richardsi*) on Tugidak Island, Alaska. Marine Mammal Science 17:585–600.

Laidre, K. L., J. A. Estes, M. T. Tinker, J. Bodkin, D. Monson, and K. Schneider. 2006. Patterns of growth and body condition in sea otters from the Aleutian archipelago before and after the recent population decline. Journal of Animal Ecology 75:978–989.

Loughlin, T. R., and A. E. York. 2002. An accounting of the sources of Steller sea lion mortality. Pages 9–13 *in* D. DeMaster and S. Atkinson, editors. Steller sea lion decline: Is it food II? University of Alaska Sea Grant College Program, Fairbanks, Alaska, USA.

Mangel, M. 2006. The theoretical biologist's toolbox: quantitative methods for ecology and evolutionary biology. Cambridge University Press, Cambridge, UK.

Mangel, M., and C. W. Clark. 1988. Dynamic modeling in behavioral ecology. Princeton University Press, Princeton, New Jersey, USA.

Mangel, M., and N. Wolf. 2006. Predator diet breadth and prey population dynamics: Mechanism and modeling. Pages 277–283 *in* J. Estes, editor. Whales, whaling and ocean ecosystems. University of California Press, Berkeley, California, USA.

Martz, H. F., and R. A. Waller. 1982. Bayesian reliability analysis. John Wiley, New York, New York, USA.

Mathews, E. A., and G. W. Pendleton. 2000. Declining trends in harbor seal (*Phoca vitulina richardsi*) numbers at glacial ice and terrestrial haulouts in Glacier Bay National Park, 1992–1998. Final report to Glacier Bay National Park and Preserve, Resource Management Division, Gustavus, Alaska 99826. Cooperative agreement 9910-97-0026.

Mathworks. 2004. Matlab. Natick, Massachusetts, USA. ⟨http://www.mathworks.com/⟩

Matkin, C. O., L. G. Barrett-Lennard, and G. Ellis. 2002. Killer whales and predation on Steller sea lions. Pages 61–66 *in* D. DeMaster and S. Atkinson, editors. Steller sea lion decline: Is it food II? University of Alaska Sea Grant College Program, Fairbanks, Alaska, USA.

Matkin, C. O., L. G. Barret-Lennard, H. Yurk, D. Ellifrit, and A. W. Trites. 2007. Ecotypic variation and predatory behavior among killer whales (*Orcinus orca*) off the eastern Aleutian Islands, Alaska. Fishery Bulletin 105:74–87.

Merrick, R. L. 1997. Current and historical roles of apex predators in the Bering Sea ecosystem. Journal of Northwest Atlantic Fishery Science 22:343–355.

Merrick, R. L., M. K. Chumbley, and G. V. Byrd. 1997. Diet diversity of Steller sea lions (*Eumetopias jubatus*) and their population decline in Alaska: a potential relationship. Canadian Journal of Fisheries and Aquatic Sciences 54: 1342–1248.

Merrick, R. L., and T. R. Loughlin. 1997. Foraging behavior of adult female and young-of-the-year Steller sea lions in Alaskan waters. Canadian Journal of Zoology 75:776–786.

Merrick, R. L., T. R. Loughlin, and D. G. Calkins. 1987. Decline in abundance of the northern sea lion, *Eumetopias jubatus*, in Alaska, 1956-86. Fishery Bulletin 86:351–365.

Milette, L. L., and A. Trites. 2003. Maternal attendance patterns of Steller sea lions (*Eumetopias jubatus*) from stable and declining populations in Alaska. Canadian Journal of Zoology 81:340–348.

Norris, K. 2004. Managing threatened species: the ecological toolbox, evolutionary theory and declining-population paradigm. Journal of Applied Ecology 41:413–426.

NRC [National Research Council]. 2003. Decline of the Steller sea lion in Alaskan waters: untangling food webs and fishing nets. National Academy of Sciences Press, Washington, D.C., USA.

Parma, A. M., P. Amarasekare, M. Mangel, J. Moore, W. W. Murdoch, E. Noonburg, M. A. Pascual, H. P. Possingham, K. Shea, C. Wilcox, and D. Yu. 1998. What can adaptive management do for our fish, forests, food, and biodiversity? Integrative Biology 1:16–26.

Pascual, M. A., and M. D. Adkison. 1994. The decline of the Steller sea lion in the northeast Pacific: demography, harvest, or environment? Ecological Applications 4:393–403.

Perez, M. A., and T. R. Loughlin. 1991. Incidental catch of marine mammals by foreign and joint venture trawl vessels in the U.S. EEZ of the North Pacific, 1973-88. U.S. Department of Commerce, National Oceanographic and Atmospheric Administration, National Marine Fisheries Service, Seattle, Washington, USA.

Pitcher, K. W. 1990. Major decline in the number of harbor seals, *Phoca vitulina richardsi*, on Tugidak Island, Gulf of Alaska. Marine Mammal Science 6:121–134.

Pitcher, K. W., V. N. Burkanov, D. G. Calkins, B. J. Le Boeuf, E. G. Mamaev, R. L. Merrick, and G. W. Pendleton. 2002. Spatial and temporal variation in the timing of births of Steller sea lions. Journal of Mammalogy 82:1047–1053.

Pitcher, K. W., D. G. Calkins, and G. W. Pendleton. 1998. Reproductive performance of female Steller sea lions: an energetics-based reproductive strategy? Canadian Journal of Zoology 76:2075–2083.

Raum-Suryan, K. L., K. W. Pitcher, D. G. Calkins, J. L. Sease, and T. R. Loughlin. 2002. Dispersal, rookery fidelity, and metapopulation structure of Steller sea lions (*Eumetopias jubatus*) in an increasing and decreasing population in Alaska. Marine Mammal Science 18:746–764.

Rea, L. D., D. A. S. Rosen, and A. W. Trites. 2000. Metabolic response to fasting in 6-week-old Steller sea lion pups (*Eumetopias jubatus*). Canadian Journal of Zoology 78:890–894.

Rehberg, M. J., K. L. Raum-Suryan, K. W. Pitcher, and T. S. Gelatt. 2001. Development of juvenile Steller sea lion (*Eumetopias jubatus*) diving behavior in Alaska. Page 177 *in* 14th Biennial Conference on the Biology of Marine Mammals. Vancouver, British Columbia, Canada.

Rosen, D. A. S., and A. W. Trites. 2000. Pollock and the decline of Steller sea lions: testing the junk-food hypothesis. Canadian Journal of Zoology 78:1243–1250.

Rosen, D. A. S., and A. W. Trites. 2002. What is it about food? Examining possible mechanisms with captive Steller sea lions. Pages 45–48 *in* D. DeMaster and S. Atkinson, editors. Steller

sea lion decline: Is it food II? University of Alaska Sea Grant College Program, Fairbanks, Alaska, USA.

Rosen, D. A. S., and A. W. Trites. 2004. Satiation and compensation for short-term changes in food quality and availability in young Steller sea lions (*Eumetopias jubatus*). Canadian Journal of Zoology 82:1061–1069.

Sandegren, F. E. 1970. Breeding and maternal behavior of the Steller sea lion (*Eumetopias jubatus*) in Alaska. Thesis. University of Alaska.

Schulman, L. S. 1981. Techniques and applications of path integration. John Wiley and Sons, New York, New York, USA.

Sease, J. L., and T. R. Loughlin. 1999. Aerial and land-based surveys of Steller sea lions (*Eumetopias jubatus*) in Alaska, June and July 1997 and 1998. U.S. Department of Commerce, National Oceanographic and Atmospheric Administration, Technical Memo NMFS-AFSC-100.

Sinclair, E. H., and T. K. Zeppelin. 2002. Seasonal and spatial differences in the diet in the western stock of Steller sea lions (*Eumetopias jubatus*). Journal of Mammalogy 83:973–990.

Small, R. J., G. W. Pendleton, and K. W. Pitcher. 2003. Trends in abundance of Alaska harbor seals, 1983–2001. Marine Mammal Science 19:344–362.

Springer, A. M., J. A. Estes, G. B. Van Vliet, T. M. Williams, D. F. Doak, E. M. Danner, K. A. Forney, and B. Pfister. 2003. Sequential megafaunal collapse in the North Pacific Ocean: An ongoing legacy of industrial whaling? Proceedings of the National Academy of Science (USA) 100:12223–12228.

Stephens, D. W., and J. R. Krebs. 1986. Foraging theory. Princeton University Press, Princeton, New Jersey, USA.

Trites, A. W., and C. P. Donnelly. 2003. The decline of Steller sea lions *Eumetopias jubatus* in Alaska: a review of the nutritional stress hypothesis. Mammal Review 33:3–28.

Trites, A. W., and P. A. Larkin. 1996. Changes in the abundance of Steller sea lions (*Eumetopias jubatus*) in Alaska from 1956 to 1992: How many were there? Aquatic Mammals 22:153–166.

Trites, A. W., and B. T. Porter. 2002. Attendance patterns of Steller sea lions (*Eumetopias jubatus*) and their young during winter. Journal of Zoology, London 256:547–556.

Wade, P., et al. 2003. Commercial whaling and "whale killers": a reanalysis of evidence for sequential megafaunal collapse in the North Pacific. Fifteenth Biennial Conference on the Biology of Marine Mammals, 14–19 December 2003. Society for Marine Mammalogy Greensboro, North Carolina, USA.

Williams, T. M., J. A. Estes, D. F. Doak, and A. M. Springer. 2004. Killer appetites: assessing the role of predators in ecological communities. Ecology 85:3373–3384.

Withrow, D. E., J. C. Cesarone, L. Hiruki-Raring, and J. L. Bengtson. 2002. Abundance and distribution of harbor seals (*Phoca vitulina*) in the Gulf of Alaska (including the south side of the Alaska Peninsula, Kodiak Island, Cook Inlet and Prince William Sound) during 2001. *In* A. L. Lopez and S. E. Moore, editors. Marine Mammal Protection Act and Endangered Species Act Implementation Program 2001. AFSC Processed Report 2002-06. Alaska Fisheries Science Center, National Marine Fisheries Service, U.S. Department of Commerce, Seattle, Washington, USA.

Withrow, D. E., J. C. Cesarone, J. K. Jansen, and J. L. Bengtson. 2000. Abundance and distribution of harbor seals (*Phoca vitulina*) along the Aleutian Islands during 1999. Pages 91–115 *in* A. L. Lopez and D. P. DeMaster, editors. Marine Mammal Protection Act and Endangered Species Act implementation program 1999. U.S. Department of Commerce, National Marine Fisheries Service, Seattle, Washington, USA.

Withrow, D. E., J. C. Cesarone, J. K. Jansen, and J. L. Bengtson. 2001. Abundance and distribution of harbor seals (*Phoca vitulina richardsi*) in Bristol Bay and along the north side of the Alaska peninsula during 2000. *In* A. L. Lopez and R. P. Angliss, editors. Marine Mammal Protection Act and Endangered Species Act Implementation Program 2000. AFSC Processed Report 2001-06. Alaska Fisheries Science Center, National Marine Fisheries Service, U.S Department of Commerce, Seattle, Washington, USA.

Wolf, N. 2006. Finding behavioral mechanisms and ecological patterns in sparse and noisy data. Dissertation. Department of Biology, University of Bergen, Bergen, Norway.

Wolf, N., and M. Mangel. 2007. Strategy, compromise, and cheating in predator–prey games. Evolutionary Ecology Research 9:1293–1304.

Wolf, N., J. Melbourne, and M. Mangel. 2006. The method of multiple hypotheses and the decline of Steller sea lions in western Alaska. Pages 275–293 *in* I. Boyd, S. Wanless, and C. J. Camphuysen, editors. Top predators in marine ecosystems: their role in monitoring and management. Cambridge University Press. Cambridge, UK.

York, A. E. 1994. The population dynamics of Northern sea lions, 1975–1985. Marine Mammal Science 10:38–51.

APPENDIX A

Data and sources (*Ecological Archives* A018-070-A1).

APPENDIX B

Characterizing beta-binomial observation error in population data (*Ecological Archives* A018-070-A2).

Ecological Archives A018-070-A1

Nicholas Wolf and Marc Mangel. 2008. Multiple hypothesis testing and the declining-population paradigm in Stellar sea lions. *Ecological Applications* 18:1932–1955.

Appendix A. Data and sources.

TABLE A1. Steller sea lion rookery locations, in order from west to east. The two-letter codes are used to identify each rookery in the space-time plots. The first 34 rookeries are in the western population. The last three are part of the eastern population, in SE Alaska, and their data were not used in this study.

	Rookery (W to E)	North Latitude	West Longitude	Years of SSL census data
AW	Attu/Cape Wrangell	52.91667	187.54083	98, 02
AG	Agattu	52.38567	186.4695	79, 88-90
KC	Kiska/Cape St. Stephen	51.88334	182.79416	79, 89, 90, 94, 98, 02
KL	Kiska/Lief Cove	51.95333	182.65884	79, 85, 89, 90, 94, 98, 02
AY	Ayugadak	51.756	181.595	79, 85, 90, 94, 98, 02
AM	Amchitka/Column Rock	51.53867	181.17867	79, 90, 94, 98, 02
SP	Semisopochnoi/Pochnoi	51.955	180.23334	94, 98
UH	Ulak/Hasgox Point	51.31334	178.98752	79, 85, 90, 94, 98, 02
TA	Tag	51.55833	178.575	85, 90, 94, 98, 02
GR	Gramp Rock	51.48117	178.343	85, 90, 94, 98, 02
AD	Adak	51.6075	176.9725	85, 90, 94, 98, 02
KN	Kasatochi/North Point	52.18517	175.51666	79, 85, 90, 94, 98, 02
AS	Amlia/Sviech. Harbor	52.03	173.39833	90, 98
AK	Agligadak	52.1015	172.90384	90, 98
SS	Seguam/Saddleridge	52.35058	172.56667	79, 85, 89, 90, 94, 98
YU	Yunaska	52.69	170.60583	79, 85, 90, 94, 98, 00, 02
AU	Adugak	52.91167	169.175	85, 90, 94, 98, 00, 02
OG	Ogchul	52.99517	168.40402	85, 94, 98
BF	Bogoslof/Fire Island	53.92822	168.03416	73, 79, 85, 89-91, 98, 00, 02
AC	Akutan/Cape Morgan	54.05906	166.02777	85, 90, 92, 94, 98
AB	Akun/Billings head	54.29321	165.53142	85, 90, 91, 94, 98, 00, 02
UC	Ugamak Complex	54.21767	164.78799	85, 86, 89-91, 96-98
SL	Sea Lion Rock (Amak)	55.46367	163.20166	98, 02
CR	Clubbing Rocks	54.706	162.4455	78, 79, 92, 94, 98
PR	Pinnacle Rock	54.76758	161.76422	78, 79, 91, 94, 98, 00, 02
CH	Chernabura	54.75867	159.57278	78, 79, 86, 90, 92, 94, 98
AT	Atkins	55.05333	159.28999	78, 79, 86, 90, 91, 94, 96, 98, 00, 02
CT	Chowiet	56.007	156.69183	78, 79, 89, 90, 92, 94, 98

CF	Chirikof	55.7745	155.69133	78, 79, 89-91, 94, 98, 00, 02
SU	Sugarloaf	58.8875	152.03999	76, 78, 79, 89, 90, 92, 94, 97, 98
MA	Marmot	58.19625	151.83176	78, 79, 86, 88, 89, 91, 94, 96-98
OP	Outer (Pye)	58.84583	150.39583	76, 78, 79, 89-91, 94, 97, 98, 00, 02
WF	Wooded (Fish)	59.88167	147.34416	94, 96, 98, 00
SR	Seal Rocks	60.163	146.83833	76, 78, 79, 89, 90, 91, 94, 96, 98
WS	White Sisters	57.635	136.25667	90-92, 94, 96, 98
HZ	Hazy	55.86666	134.56667	79, 89, 90, 91, 94, 96, 98
FC	Forrester Complex	54.83833	133.52667	73, 79, 82, 90, 91

TABLE A2. Major prey species of Steller sea lions (based on Sinclair and Zeppelin 2002).

Common name	**Species name(s)**
Walleye Pollock	*Theragra chalcogramma*
Atka Mackerel	*Pleurogrammus monopterygius*
Pacific Salmon	*Oncorhynchus spp.*
Pacific Cod	*Gadus macrocephalus*
Arrowtooth Flounder	*Atheresthes stomias*
Pacific Herring	*Clupea pallasi*
Pacific Sand Lance	*Ammodytes hexapterus*
Irish Lords	*Hemilepidotus spp.*
Cephalopods	Class *Cephalopoda*
Capelin	*Mallotus villosus*
Rockfishes	Family *Scorpaenidae*

TABLE A3. Parameters and data sources used in the model.

Parameter	Value	Description and (Source)
$J(i,t)$	See Fig. 2a	Number of pups *counted* at rookery i in year t
$N_{obs}(i,t)$	See Fig. 2b	Number of non-pups *observed* at rookery i in year t (Alaska Fisheries Science Center/National Marine Mammal Lab)
α	6	1st parameter of beta-binomial observation error distribution
β	2	2nd parameter of beta-binomial observation error distribution (Est. from repeated counts at one rookery; see Appendix B)
ρ_0	0.776	Pre-decline estimate of pup recruitment rate
σ_0	0.858	Pre-decline estimate of non-pup annual survival rate
ϕ_0	0.197	Pre-decline estimate of fecundity rate (Based on Calkins and Pitcher 1982, Holmes and York 2003)
S_2	7 days	Maximum fasting period for pups before starvation
S_3	21 days	Maximum fasting period for female non-pups before starvation (Chosen to exceed durations of observed fasts; see text)
S_1	14 days	Maximum fasting period before termination of pregnancy (Intermediate between S_2 and S_3; see text for details)
r_{forage}	300 km	Radius of foraging zone in which prey and HS densities affect SSL (Following Gerber and van Blaricom 2001)
$r_{fisheries}$	20 km	Radius of interaction zone in which fisheries activities impact SSL (Following Hennen 2006)
$\lambda_{prey}(i,t)$	See Fig. 4a	Prey density in foraging zone of rookery i/year t, excluding pollock
$\lambda_{pollock}(i,t)$	See Fig. 4b	Pollock density in foraging zone of rookery i in year t (National Marine Fisheries Service groundfish survey trawl data)
$\mu(i,t)$	See Fig. 4c	Number of fishing events in interaction zone of rookery i in year t (NMFS fishery observer data compiled by Daniel Hennen)
$h(i,t)$	See Fig. 4d	Harbor seal density within 300 km of rookery i in year t (Compiled from various sources; see text)
h_{crit}	4500	HS density separating rookeries w/ rising and falling SSL trends (See text, Figure 5, and Mangel and Wolf 2006)

LITERATURE CITED

Calkins, D., and K. W. Pitcher. 1982. Population assessment, ecology and trophic relationships of Steller sea lions in the Gulf of Alaska. Pages 447–546 *in* Environmental assessment of the Alaskan continental shelf, Final report. U.S. Department of Commerce and U.S. Department of the Interior. Washington,

D.C., USA.

Gerber, L. R., and G. R. Van Blaricom. 2001. Implications of three viability models for the conservation status of the western population of Steller sea lions (Eumetopias jubatus). Biological Conservation 102:261–269.

Hennen, D. 2006. Associations between the Alaska Steller sea lion decline and commercial fisheries. Ecological Applications 16:704–717.

Holmes, E. E., and A. E. York. 2003. Using age structure to detect impacts on threatened populations: A case study with Steller sea lions. Conservation Biology 17(6):1794–1806.

Mangel, M., and N. Wolf. 2006. Predator diet breadth and prey population dynamics: Mechanism and modeling. Pages 277–283 *in* J. Estes, editor. Whales, whaling and ocean ecosystems. University of California Press. Berkeley, Californai, USA.

Sinclair, E. H., and T. K. Zeppelin. 2002. Seasonal and spatial differences in the diet in the western stock of Steller sea lions (Eumetopias jubatus). Journal of Mammalogy 83:973–990.

[Back to A018-070]

Ecological Archives **A018-070-A2**

Nicholas Wolf and Marc Mangel. 2008. Multiple hypothesis testing and the declining-population paradigm in Stellar sea lions. ***Ecological Applications*** 18: 1932–1955.

Appendix B. Characterizing beta-binomial observation error in population data.

Abstract

Researchers are often called upon to make inferences about the size of a population for which the only available data are undercounts. The solution requires simultaneous estimation of population size and parameters of the error structure, using the beta-binomial distribution (or similar) to accommodate variation in the sighting probability between counts. We show how to do this, and explain how to solve the related problem of characterizing the beta-binomial error using numerical integration, Bayes's theorem, and uniform priors. We test the methods by estimating the parameters used to create 100 simulated beta-binomial datasets. The resultant 95% confidence regions include the respective "true" values in 94 of the 100 trials – very close to the expected 95 out of 100. Census data from the western population of Steller sea lions provide a real-world application.

Introduction

When organisms are mobile or cryptic, so that some are inevitably missed in any systematic survey, assessment of population size becomes a problem of simultaneously estimating the true number of individuals and the probability of counting them. This scenario applies equally to pinnipeds counted in aerial photographs of a rookery, concealed rare plants counted along a standardized search path, disease incidence estimated using questionnaires or hospital entry rates, or any other survey where each individual is counted once or not at all, and the result is an undercount. The underlying observation error structure is the same in each case (Slade et al. 2003).

Much work has gone into understanding the variation of sighting probability among individuals (for example, see Everson and Bradlow 2002, Gardner and Mangel 1996). In contrast, we are concerned here with the variation of mean sighting probability between census events. Our methods involve direct calculation of Bayesian posterior probabilities (or linear multiples thereof) using numerical integration and explicit uniform (uninformative) prior probability distributions.

Observation error paradigms: Binomial and Beta-binomial

The traditional solutions for estimating sighting probability (re-sighting or removal approaches) are often unavailable, particularly in the case of historical data, and we are left to estimate the parameters of the observation error distribution directly from the census data. It is important in such cases to consider the mechanisms that prevent some individuals from being counted. For example, the sighting probability for pinnipeds within the frame of a photograph is essentially 1, and animals are missed because they were not on the rookery at the time of the photo. With rare plants, individuals are missed because they are simply difficult to see. In the case of disease surveys, sick individuals may be asymptomatic and/or unaware that they are infected, or they may be aware but reluctant to reveal themselves for a variety of reasons. The consequence is underestimation of the number infected.

In all cases, if individuals are exchangeable (that is, if we can assume that all individuals are equally likely to be missed), then the distribution of observed counts is binomial given the unknown true population, N_{true}, and the unknown individual sighting probability, P:

$$\Pr\left\{N_{obs} = k \middle| N_{true} = n, P = p\right\} = \binom{n}{k} p^k (1-p)^{n-k} \quad \text{(B.1)}$$

The binomial expected mean and variance are

$$E\left\{N_{obs}\right\} = N_{true} P \approx \overline{N}_{obs} \qquad \text{(for multiple counts)} \quad \text{(B.2)}$$

$$E\left\{V_{obs}\right\} = N_{true} P(1-P) \quad \text{(B.3)}$$

Thus, the binomial implies a specific relationship between the expected mean and variance of the data.

$$\frac{E\left\{V_{obs}\right\}}{E\left\{N_{obs}\right\}} = (1-P) \quad \text{(B.4)}$$

Because (1-P) is less than or equal to 1, the expected variance in binomial data is less than or equal to the expected mean.

When the probability of detection varies between counts, the variance quickly exceeds the mean, resulting in overdispersed data and necessitating a more sophisticated model to characterize the variation in P. This variation comes from a variety of sources. In the pinniped example, the fraction of individuals hauled out at the time of the aerial photo varies due to fluctuations in weather, tidal heights, and foraging conditions. In the case of rare plant surveys, variation in P among counts might arise from changing light levels or from differences in the success rates of different observers (if different observers perform the different censuses). Sometimes these effects can be estimated using covariate analysis (e.g., Mathews and Pendleton

2000, Ver Hoef and Frost 2003, Hayward et al. 2005), but considerable unexplained variation in P generally remains.

The problem of simultaneously estimating the number of trials and the distribution of the probability of success in a binomial process is an exceptionally difficult one (Hilborn and Mangel 1997). The best existing solution is the beta-binomial model (Martz and Waller 1982; Evans et al 2000), in which the sighting probability is assumed to follow a beta probability density and, conditioned on that, the number of individuals observed is binomially distributed with unknown true total number and beta-distributed sighting probability.

The Beta-binomial distribution has three parameters: N_{true} is the true population size, as above; α and β define the Beta distribution of P among censuses. (Note: We use the following shorthand to denote the probability densities of continuous variables.

$$\Pr\{X \approx x\} = \Pr\{x \le X \le x + dx\})$$

$$\Pr\{P \approx p\} = p^{\alpha-1}(1-p)^{\beta-1}\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}dp \quad \text{with} \quad \alpha, \beta > 1 \tag{B.5}$$

where $\Gamma(x)$ is the gamma function (Abramowitz and Stegun 1974) satisfying the recursion

$$\Gamma(x+1) = x\Gamma(x). \tag{B.6}$$

The mean of the Beta distribution, or the expectation of P, is

$$E\{P\} = \frac{\alpha}{\alpha+\beta}. \tag{B.7}$$

The probability of a particular count given α, β, and N_{true}, or the likelihood of α, β, and N_{true} given the count (Hilborn and Mangel 1997) is calculated by integrating the binomial across the distribution of P. The result is an "integrated likelihood function" (Kalbfleisch and Sprott 1970):

$$\Pr\left\{N_{obs} = k \middle| \alpha, \beta, N_{true}\right\} = \int_{p=0}^{1} p^{\alpha-1}(1-p)^{\beta-1} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \binom{N_{true}}{k} p^{k}(1-p)^{N_{true}-k}\, dp \qquad \text{(B.8)}$$

or, equivalently,

$$\Pr\left\{N_{obs} = k \middle| \alpha, \beta, N_{true}\right\} = \binom{N_{true}}{k} \frac{\Gamma(\alpha+k)\Gamma(\beta+N_{true}-k)\Gamma(\alpha+\beta)}{\Gamma(\alpha+\beta+N_{true})\Gamma(\alpha)\Gamma(\beta)}. \qquad \text{(B.9)}$$

The likelihood of all observed counts (k_1, k_2, $k_3 \ldots k_j$) given a particular set of parameter values is the product of the likelihoods of all observed data:

$$\begin{aligned} &Likelihood\left\{\alpha, \beta, N_{true} \middle| data\right\} = \\ &\Pr\left\{data = k_1, k_2, k_3 \ldots k_j \middle| \alpha, \beta, N_{true}\right\} = \prod_{i=1}^{j} \binom{N_{true}}{k_i} \frac{\Gamma(\alpha+k_i)\Gamma(\beta+N_{true}-k_i)\Gamma(\alpha+\beta)}{\Gamma(\alpha+\beta+N_{true})\Gamma(\alpha)\Gamma(\beta)} \end{aligned} \qquad \text{(B.10)}$$

Estimating α and β, with N_{true} as a nuisance parameter

When α, β, and N_{true} are unknown but we are only interested in α and β (for example, when we need to characterize observation error for a population model), the most direct approach is to convert the likelihood into a posterior probability and calculate its expectation across the range of values that might be taken by the "nuisance parameter," N_{true}.

Bayes's theorem (Hilborn and Mangel 1997) provides the bridge between likelihood and posterior probability, given a set of prior probability distributions for the parameters (denoted here as $\Pr_{prior}\{\alpha, \beta\}$). Given N_{true}, the posterior probability density of α and β is

$$\Pr\{\alpha,\beta|data,N_{true}\}=\frac{\Pr\{data|\alpha,\beta,N_{true}\}\Pr_{prior}\{\alpha,\beta|N_{true}\}}{\Pr\{data|N_{true}\}}$$
$$=\frac{\Pr\{data|\alpha,\beta,N_{true}\}\Pr_{prior}\{\alpha,\beta|N_{true}\}}{\int_{a=1}^{\infty}\int_{b=1}^{\infty}\Pr_{prior}\{\alpha\approx a,\beta\approx b|N_{true}\}\Pr\{data|\alpha=a,\beta=b,N_{true}\}db\ da} \quad \text{(B.11)}$$

Assuming uniform priors for α and β, this reduces to:

$$\Pr\{\alpha,\beta|data,N_{true}\}=\frac{\Pr\{data|\alpha,\beta,N_{true}\}}{\int_{a=1}^{\infty}\int_{b=1}^{\infty}\Pr\{data|\alpha=a,\beta=b,N_{true}\}db\ da} \quad \text{(B.12)}$$

And finally, substituting in the terms from Eq. B.10:

$$\Pr\{\alpha,\beta|data,N_{true}\}=\frac{\prod_{i=1}^{j}\binom{N_{true}}{k_i}\frac{\Gamma(\alpha+k_i)\Gamma(\beta+N_{true}-k_i)\Gamma(\alpha+\beta)}{\Gamma(\alpha+\beta+N_{true})\Gamma(\alpha)\Gamma(\beta)}}{\int_{a=1}^{\infty}\int_{b=1}^{\infty}\prod_{i=1}^{j}\binom{N_{true}}{k_i}\frac{\Gamma(a+k_i)\Gamma(b+N_{true}-k_i)\Gamma(a+b)}{\Gamma(a+b+N_{true})\Gamma(a)\Gamma(b)}db\ da} \quad \text{(B.13)}$$

The term in the denominator is the likelihood sum across an α/β plane conditioned on the value of N_{true}. Equation B.13 therefore amounts to normalizing the likelihood within each α/β plane.

After canceling out the binomial coefficient, Equation B.13 reduces to:

$$\Pr\{\alpha,\beta|data,N_{true}\}=\frac{\prod_{i=1}^{j}\frac{\Gamma(\alpha+k_i)\Gamma(\beta+N_{true}-k_i)\Gamma(\alpha+\beta)}{\Gamma(\alpha+\beta+N_{true})\Gamma(\alpha)\Gamma(\beta)}}{\int_{a=1}^{\infty}\int_{b=1}^{\infty}\prod_{i=1}^{j}\frac{\Gamma(a+k_i)\Gamma(b+N_{true}-k_i)\Gamma(a+b)}{\Gamma(a+b+N_{true})\Gamma(a)\Gamma(b)}db\ da} \quad \text{(B.14)}$$

The next step is to remove the conditioning on the "nuisance parameter" (N_{true} in this case) and calculate the posterior probability for α and β given only the data (Mangel and Clark 1988). To do this, we add up the posterior probabilities corresponding to all plausible values of N_{true}, weighting each one by the prior probability density, $\Pr_{prior}\{N_{true} = n\}$.

$$\begin{aligned}\Pr\{\alpha,\beta|data\} &= \sum_{all\ n}\Pr_{prior}\{N_{true}=n\}\Pr\{\alpha,\beta|data,N_{true}=n\} \\ &= \sum_{n=\max(k)}^{\infty}\Pr_{prior}\{N_{true}=n\}\frac{\prod_{i=1}^{j}\frac{\Gamma(\alpha+k_i)\Gamma(\beta+n-k_i)\Gamma(\alpha+\beta)}{\Gamma(\alpha+\beta+n)\Gamma(\alpha)\Gamma(\beta)}}{\int_{a=1}^{\infty}\int_{b=1}^{\infty}\prod_{i=1}^{j}\frac{\Gamma(a+k_i)\Gamma(b+n-k_i)\Gamma(a+b)}{\Gamma(a+b+n)\Gamma(a)\Gamma(b)}db\ da}\end{aligned} \quad \text{(B.15)}$$

This is the Bayesian posterior for α and β given only the data.

The maximum posterior probability estimates for α and β are the coordinates of the point at which Equation B.15 is maximized. A 95% Bayesian confidence region around that point can be constructed by finding the contour that surrounds 95% of the total posterior probability mass. That total is calculated using numerical integration across the region of α/β space in which likelihood is appreciable (Lee and Lio 1999). See the Supplement for suggestions regarding the choice of upper limits.

Model Validation using Simulated Data

In order to test the model's accuracy, we produced 100 simulated beta-binomial datasets and then estimated α and β according to the methods described above. Each dataset consisted of 20 simulated counts produced using values for α, β, and N_{true} drawn from uniform random distributions with ranges of [1, 20], [1, 20], and [1, 200], respectively. Figure B.1 shows the percentage of trials in which the posterior likelihood percentile of the true α and β values fell below any given quantile. (If the model is appropriate and the priors are correctly specified, the true values should fall within the n^{th} percentile contour in n% of the trials, for any given n).

The true values fell within the 95th percentile of the posterior likelihood surface in 94 cases out of 100, which is very close to the expected 95 out of 100. However, for lower percentiles, the model appears to be more accurate than necessary. This is probably an artifact of the fact that the likelihood was calculated by taking the expectation across a broad range of N_{true} values, while the true range was limited to [1, 200].

Example: Steller sea lions

Eight Steller sea lion counts, from replicate aerial photos of a single rookery in a single month, provide a practical application of the approach. Data are from Outer (Pye) Island, June 1992, from the National Marine Mammal Laboratory (http://nmml.afsc.noaa.gov/AlaskaEcosystems/sslhome/stellerhome.html).

Day	10	11	13	14	15	17	23	26
Count	369	391	371	370	477	242	481	319

The variance (5343.5) greatly exceeds the mean (377.5), indicating that there is variation in P. The peak of the posterior probability (Eq. B.14) drops from a high of 0.87 in the α/β plane corresponding to N_{true} = 481, to less than 0.1% of this value in the plane corresponding to N_{true} = 2845 (Fig. B.1). Therefore, we sum over the range from 481 to 2845. Figure B.2 shows the resultant posterior probability surface, with a peak near $\alpha = 6$, $\beta = 2$. The 95% confidence region is large, but the peak is relatively well defined.

These point estimates of α and β produce the Beta distribution shown in Fig. B.3. The mean sighting probability is 0.75, which agrees well with subjective estimates of the mean sighting probability produced by experts in the field (Tim Ragen, *personal communication*).

A single observation (N_{obs} = 406) was made at this rookery in June of 1994, providing an opportunity to describe the probability distribution for N_{true} given N_{obs}, α, and β (Eq. B.9). This

is shown in Fig. B.4. The mode of the distribution is 473 and the expectation is approximately 609. The 95% confidence interval, defined as the range around the mode containing 95% of the total area under the curve, is [407, 980].

Conclusions

Our method should provide unbiased estimates of the Beta-binomial parameters, even with small datasets, although the 95% confidence region can be quite large. The Steller sea lion example - the first real-world test of the approach - appears to produce reasonable estimates. We suggest that this approach is applicable to any system in which each individual is counted once or not at all, and the result is an undercount.

Literature cited

Abramowitz, M., and I. A. Stegun. 1964. Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables. Dover, New York, New York, USA.

Evans, M., N. Hastings, and B. Peacock. 2000. Statistical Distributions, Third edition. Wiley, New York, New York, USA.

Everson, P. J., and E. T. Bradlow. 2002. Bayesian Inference for the Beta-Binomial Distribution via Polynomial Expansions. Journal of Computational and Graphical Statistics 11:202-207.

Gardner, S. N., and M. Mangel. 1996. Mark-Resight Population Estimation with Imperfect Observations. Ecology 77:880-884.

Hayward, J. L., S. M. Henson, C. J. Logan, C. R. Parris, M. W. Meyer, and B. Dennis. 2005. Predicting Numbers of Hauled-out Harbour seals: A Mathematical Model. Journal of Applied Ecology 42:108-117.

Hilborn, R., and M. Mangel. 1997. The Ecological Detective: Confronting Models with Data. Princeton University Press, Princeton, New Jersey, USA.

Kalbfleisch, J. D., and D. A. Sprott. 1970. Application of Likelihood Methods to Models Involving Large Numbers of Parameters. Journal of the Royal Statistical Society. Series B (Methodological) 32:175-208.

Lee, J. C., and Y. L. Lio. 1999. A Note on Bayesian Estimation and Prediction for the Beta-Binomial Model. Journal of Statistical Computation and Simulation 63:73-91.

Mangel, M., and C. W. Clark. 1988. Dynamic Modeling in Behavioral Ecology. Princeton University Press, Princeton, New Jersey, USA.

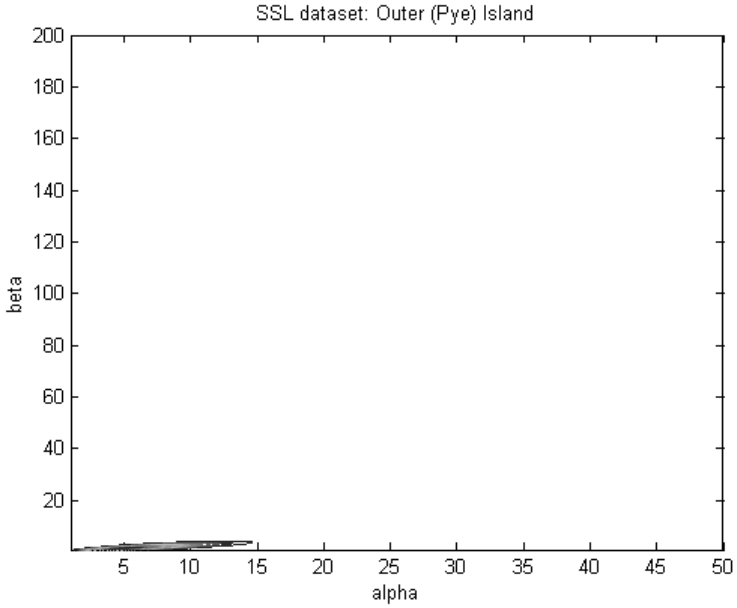
Martz, H. F., and R. A. Waller. 1982. Bayesian Reliability Analysis. John Wiley and Sons, New York, New York, USA.

Mathews, E. A., and G. W. Pendleton. 2000. Declining Trends in Harbor Seal (*Phoca vitulina richardsi*) Numbers at Glacial Ice and Terrestrial Haulouts in Glacier Bay National Park, 1992-1998. Cooperative Agreement 9910-97-0026, Glacier Bay National Park and Preserve, Resource Management Division, Gustavus, Alaska, USA.

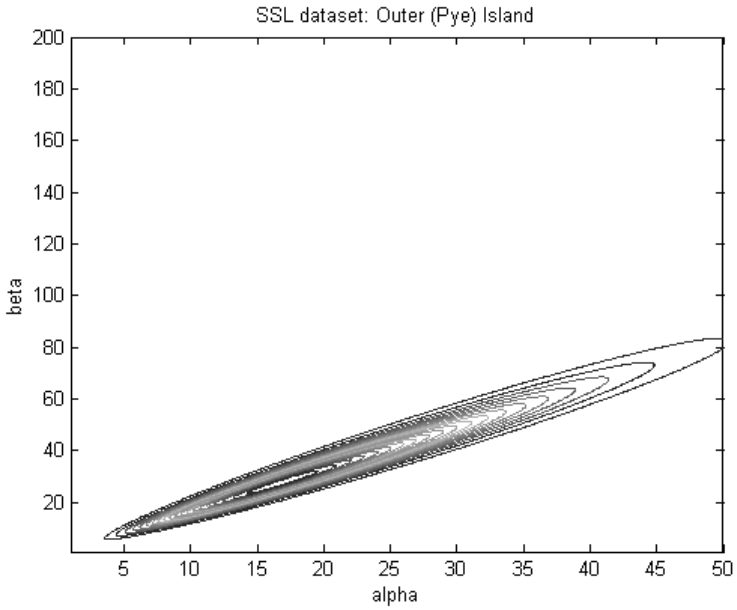
Slade, N. A., H. M. Alexander, and W. D. Kettle. 2003. Estimation of Population Size and Probabilities of Survival and Detection in Mead's Milkweed. Ecology 84:791-797.

Ver Hoef, J. M. V., and K. J. Frost. 2003. A Bayesian Hierarchical Model for Monitoring Harbor Seal Changes in Prince William Sound, Alaska. Environmental and Ecological Statistics 10:201-219.

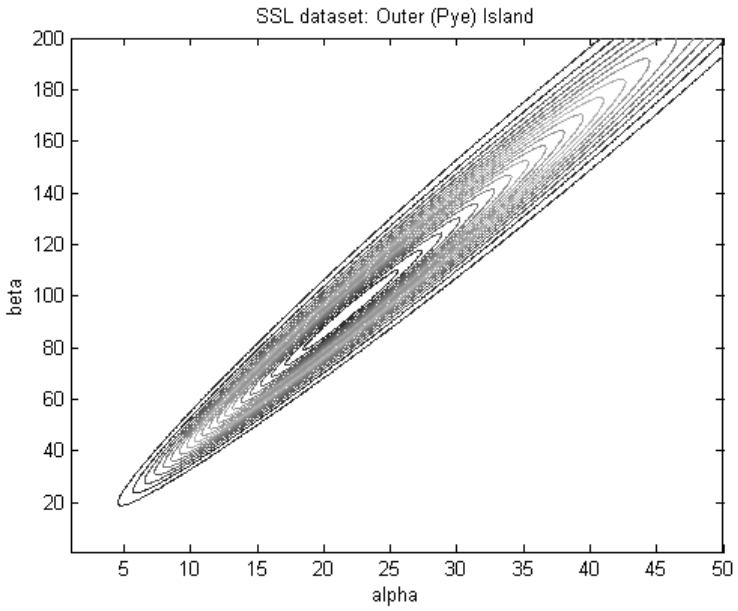
B1a: $N_{true} = 500$



B1b: $N_{true} = 1000$



B1c: $N_{true} = 2000$



B1d: $N_{true} = 4000$

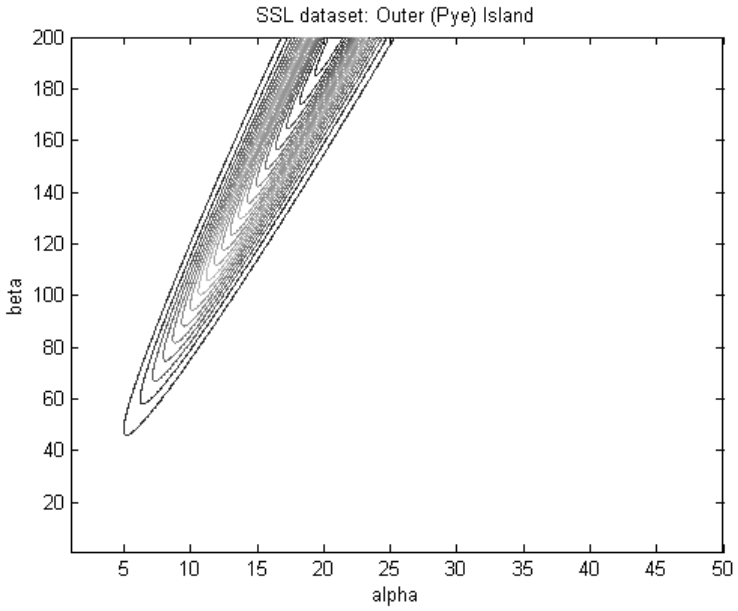
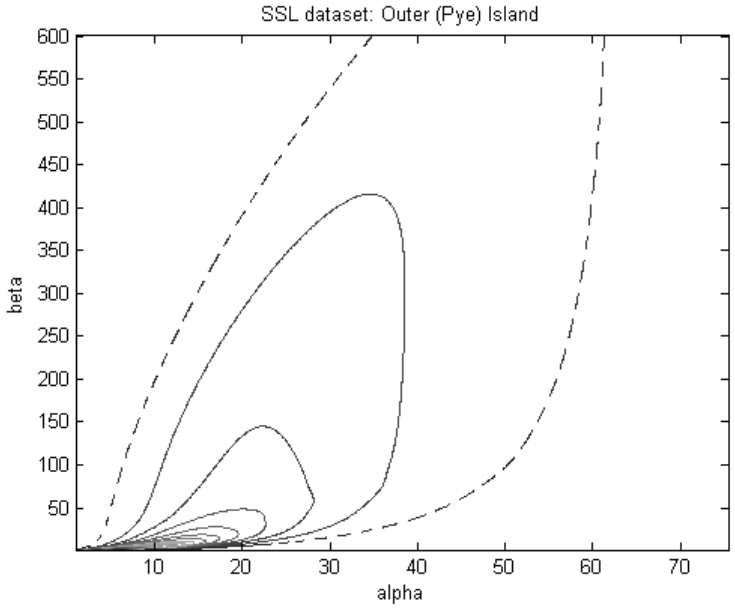


FIG. B1. Likelihood contour plots for α and β given several different values of N_{true}, calculated using the Steller sea lion data. The figures show α/β planes corresponding to $N_{true} = 500$ (a), 1000 (b), 2000 (c), and 4000 (d).

B2a



B2b

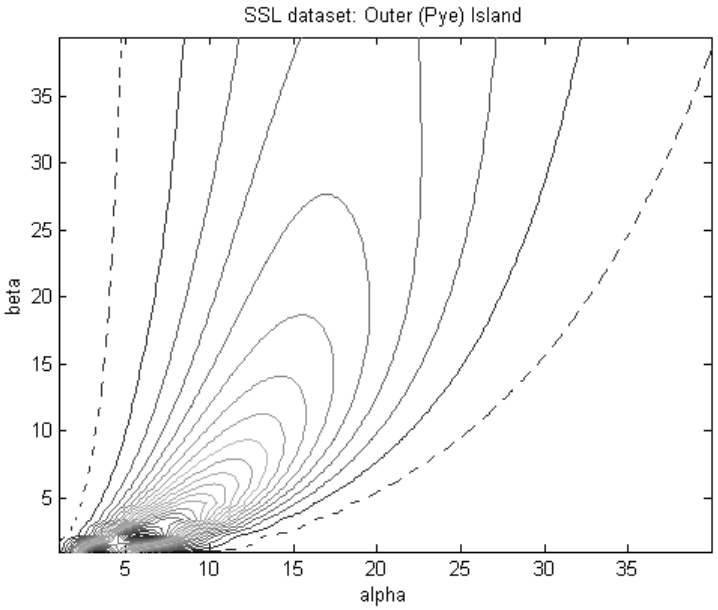


FIG. B2. Posterior probability surface for the Steller sea lion data across the range where the summed likelihood is appreciable (a), and at a finer scale (b). The dotted contour line indicates the 95% confidence region.

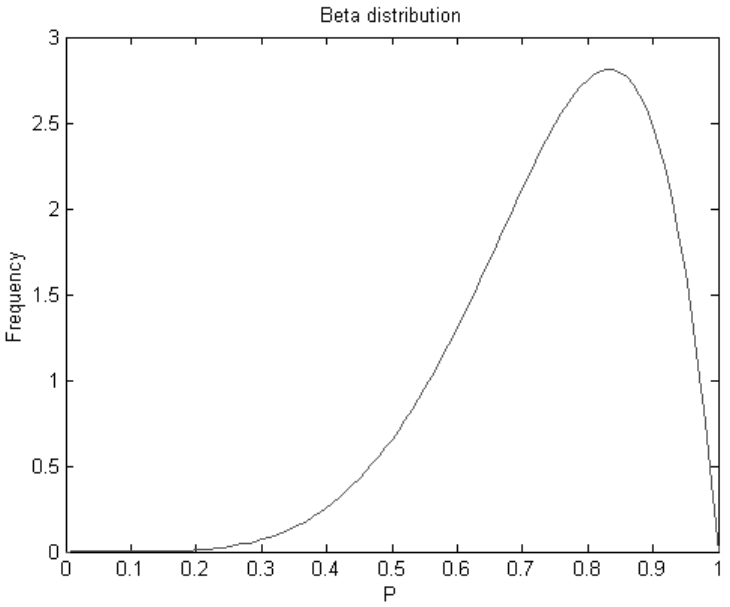


FIG. B3. Beta distribution corresponding to the Bayesian posterior mode values of $\alpha = 6$ and $\beta = 2$.

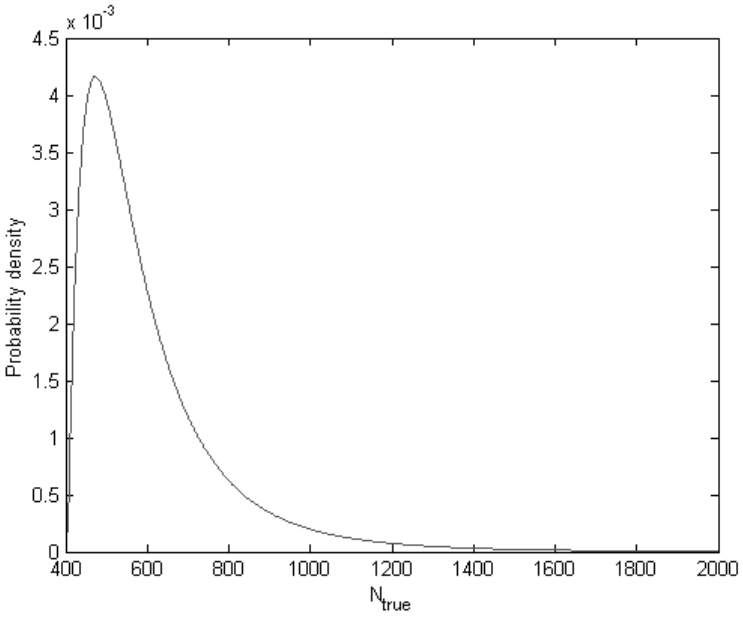


FIG. B4. Posterior probability distribution for N_{true} given a single observation, $N_{obs} = 406$, and using fixed values for α and β (that is, $\alpha = 6$ and $\beta = 2$).

Supplement: Details regarding the choice of upper limits for summations and numerical integrations.

Several simplifications are necessary in order to make the calculations tractable. The factorials are calculated as gamma densities according to $x! = \Gamma(x+1)$. The integrations across *a* and *b* (representing α and β) are approximated numerically. The upper limits of those numerical integrations, and of the summation across *n* (representing N_{true}), are chosen to include the range in which the summed likelihood is appreciable. Since the likelihood functions and posterior probability distributions described in this paper are all unimodal, the choice of limits is relatively straightforward.

Choosing the limits

Using the method of moments allows us to calculate rough estimates of the parameter values (or at least the relationships between parameter values) and narrow our search to the parameter space near those estimates, saving considerable computation time. Specifically, the mean ($\overline{N}_{obs}$) and variance (V_{obs}) of the observed data provide some information about the relationship between α, β, and N_{true}. Since we have three parameters to estimate, but only two moments with which to estimate them, we cannot solve for the parameters outright, but we can find relations for any one in terms of another, as shown by the following calculations.

The expected mean and variance of the Beta-binomial are

$$E\{\overline{N}_{obs}\} = \frac{N_{true}\alpha}{\alpha+\beta} \tag{B.16}$$

$$E\{V_{obs}\} = \frac{N_{true}\alpha\beta}{(\alpha+\beta)^2}\left(1+\frac{N_{true}-1}{1+\alpha+\beta}\right) = \frac{N_{true}\alpha\beta}{(\alpha+\beta)^2}\frac{\alpha+\beta+N_{true}}{\alpha+\beta+1} \tag{B.17}$$

We begin by setting $\overline{N}_{obs}$ equal to $E\{\overline{N}_{obs}\}$ and V equal to $E\{V_{obs}\}$, and after a bit of algebra we arrive at equations for $\hat{\alpha}$ and $\hat{\beta}$ (point estimates of α and β) in terms of N_{true}, given $\overline{N}_{obs}$ and V_{obs}.

$$\hat{\alpha} = \overline{N}_{obs} \frac{\overline{N}_{obs}\left(N_{true} - \overline{N}_{obs}\right) - V_{obs}}{N_{true}V_{obs} - \overline{N}_{obs}\left(N_{true} - \overline{N}_{obs}\right)} \tag{B.18}$$

$$\hat{\beta} = \left(N_{true} - \overline{N}_{obs}\right) \frac{\overline{N}_{obs}\left(N_{true} - \overline{N}_{obs}\right) - V_{obs}}{N_{true}V_{obs} - \overline{N}_{obs}\left(N_{true} - \overline{N}_{obs}\right)} \tag{B.19}$$

The asymptotic curve defined by Eqs. B.18 and B.19 is clearly visible in Fig. B1 as the path followed by the likelihood peak in successive α/β planes. As N_{true} increases to ∞, $\hat{\alpha}$ increases to an asymptote at $\frac{\overline{N}_{obs}^2}{V_{obs} - \overline{N}_{obs}}$ and $\hat{\beta}$ increases indefinitely.

This behavior of α and β makes it possible to set reasonable upper limits for these parameters when performing the double numerical integrations within α/β planes (given N_{true}) that appear in Eqs. B.11 through B.15: For each value of N_{true}, the likelihood is highest when α and β are close to the values predicted by Eqs. B.18 and B.19. Therefore, a straightforward way to limit the search space to a region containing the vast majority of the total probability is to sum across the rectangle defined by $\left(1 \le \alpha \le r\hat{\alpha},\ 1 \le \beta \le r\hat{\beta}\right)$, with r chosen to include at least 95% of the estimated total likelihood. In the Steller sea lion example, r=12 was more than sufficient for all values of N_{true} considered. Summation across α/β planes corresponding to different values of N_{true} requires bilinear interpolation because the α and/or β step size is smaller for lower N_{true}.

In other cases, the limits must be found iteratively by repeatedly extending the limit and calculating the likelihood or probability sum until it approaches an asymptote, or (equivalently) until the marginal value falls below some small fraction of the peak value. In the Steller sea lion example, we set the upper cutoff for N_{true} (in Eq. B.15) to the value at which the likelihood increment falls below 1% of the cumulative likelihood sum for all lower values of N_{true} across the region of interest in α/β space (with the limits of this region chosen iteratively to encompass the vast majority of the area where likelihood is appreciable).