



RESEARCH ARTICLE

Comparing reproductive success of a colonial seabird, the Magellanic Penguin, estimated by coarse- and fine-scale temporal sampling

Ginger A. Rebstock^{1,2*} and P. Dee Boersma^{1,2}

¹ Department of Biology, University of Washington, Seattle, Washington, USA

² Wildlife Conservation Society, Bronx, New York, USA

* Corresponding author: gar@u.washington.edu

Submitted July 15, 2016; Accepted January 24, 2017; Published April 5, 2017

ABSTRACT

Fecundity, an important demographic variable for wildlife populations, is time consuming and expensive to measure. For seabirds, reproductive success is often estimated from 2 surveys of the colony, 1 during incubation and 1 during chick rearing. Using 33 yr of data on Magellanic Penguins (*Spheniscus magellanicus*), we compared reproductive success estimated from 2 surveys per season, from infrequent nest checks, and from daily nest checks to answer 4 questions: (1) Is reproductive success estimated from 2 surveys per season accurate enough for population models (within 15% of the 'true' value)? We found that the answer was yes, if surveys were done on optimal dates and chick counts were adjusted by the percentage of chicks that were heavy enough to fledge. Optimal dates depended on the synchrony and phenology of the species. If penguins bred particularly late or early in a given year, estimates were wrong even if counts were done on the average optimal date. (2) Can counts of occupied nests be adjusted based solely on date? We determined no, not without additional information on egg dates and egg and chick losses. (3) Can long-term trends in reproductive success be detected from 2 surveys per season? Our answer was yes; even using biased data from counts on suboptimal dates, trends with similar slopes were detected using both methods (2 surveys vs. nest checks). (4) How often must nests be checked to get reproductive success estimates comparable with those calculated from daily nest checks? We discovered that checking nests every 2 days was as good as doing daily checks. Reproductive success was overestimated by <5% when checks were conducted every 3–8 days and by <15% (in 90% of years) when checks were done every 9–30 days. The degree of overestimation depended on the timing of nest checks relative to egg laying, loss, and hatching. We conclude that reproductive success can be estimated from 2 surveys, but only if timing, synchrony, and variability of breeding are also known.

Keywords: fecundity, Magellanic Penguin, reproductive success, *Spheniscus magellanicus*, sampling frequency, survey

Comparación del éxito reproductivo de un ave marina colonial, *Spheniscus magellanicus*, estimado mediante muestreos temporales de escala gruesa y fina

RESUMEN

La fecundidad, una variable demográfica importante en las poblaciones de vida silvestre, requiere mucho tiempo y es costosa para medir. Para las aves marinas, el éxito reproductivo es usualmente estimado a partir de dos muestreos de la colonia, uno durante la incubación y otro durante el crecimiento de los polluelos. Usando 33 años de datos de *Spheniscus magellanicus*, comparamos el éxito reproductivo estimado a partir de dos muestreos por temporada, mediante controles infrecuentes de los nidos y controles diarios de los nidos, para responder cuatro preguntas. 1) ¿Es el éxito reproductivo estimado a partir de dos muestreos por temporada suficientemente preciso para los modelos poblacionales (dentro del 15% del valor 'verdadero')? Sí, si los muestreos fueron realizados en las fechas óptimas y los conteos de los polluelos fueron ajustados por el porcentaje de polluelos lo suficientemente pesados como para dejar la colonia. Las fechas óptimas dependen de la sincronía y la fenología de la especie. Si anidamiento fue particularmente tarde o temprano, las estimaciones fueron erradas incluso si los conteos fueron hechos en la fecha óptima promedio. 2) ¿Pueden los conteos de los nidos ocupados ser ajustados basados solamente en la fecha? No, se requiere información adicional sobre la fecha de los huevos y sobre las pérdidas de huevos y polluelos. 3) ¿Pueden las tendencias de largo plazo en el éxito reproductivo ser detectadas a partir de dos muestreos por temporada? Sí, incluso usando datos sesgados a partir de conteos realizados en fechas subóptimas, se detectaron tendencias con pendientes similares usando ambos métodos. 4) ¿Cuán seguido deben ser controlados los nidos para obtener un éxito reproductivo comparable con el resultado de los controles diarios de los nidos? Controlar los nidos cada dos días fue tan bueno como los controles diarios. El éxito reproductivo fue sobreestimado por <5% para controles realizados cada 3–8 días y por <15% (en el 90% de los años) para controles realizados cada 9–30 días. El grado de sobreestimación depende del

momento del control de los nidos, en relación a la puesta, la pérdida y la eclosión de los huevos. Concluimos que el éxito reproductivo puede ser estimado a partir de dos muestreos, pero solo si el momento, la sincronía y la variabilidad de la reproducción también se conocen.

Palabras clave: éxito reproductivo, fecundidad, frecuencia de muestreo, muestreo, Pingüino Magallánico, *Spheniscus magellanicus*

INTRODUCTION

Fecundity is one of the main demographic variables determining wildlife population sizes and trends. Even in long-lived species such as seabirds, reproductive success affects population trends (Jenouvrier et al. 2009, Sandvik et al. 2012, Blight et al. 2015). Hence, reproductive success is needed for demographic models, used in the management and conservation of species, and incorporated for understanding changes in population size. The collection of reproductive success data is mandated or encouraged by international and national organizations such as the Commission for the Conservation of Antarctic Marine Living Resources (CCAMLR) Ecosystem Monitoring Program (CEMP; CCAMLR 2014), the U.S. Fish and Wildlife Service (Alaska Maritime National Wildlife Refuge 2016), and the Joint Nature Conservation Committee of the UK (Walsh et al. 1995).

A change in reproductive success can serve as an early warning of population response to environmental change. Seabirds generally have high adult survival and delay the start of reproduction for a few years to more than a decade (Weimerskirch 2002). Reproductive success is more variable among years than adult survival (Cairns 1987, Waugh et al. 2015). Because of these traits, declines in population size may lag behind changes in the environment by years. Reproductive success responds more rapidly than population size to environmental changes such as weather or prey abundance (Cairns 1987, Walsh et al. 1995, Einoder 2009, Oro 2014).

The reproductive success of seabirds is often estimated by following the fates of nests throughout a breeding season (Croxall and Prince 1979, Birkhead and Nettleship 1980, Boersma et al. 1980, CCAMLR 2014, Alaska Maritime National Wildlife Refuge 2016). Checking nests frequently from before egg laying until chicks fledge yields the best estimate of reproductive success, but is also the most time-consuming, labor-intensive, and expensive method. Hence, reproductive success of seabirds is often estimated from 2 visits to the breeding colony, 1 during egg laying or incubation and 1 during chick rearing (Walsh et al. 1995, Crawford et al. 2006, Delord et al. 2008, CCAMLR 2014, Alaska Maritime National Wildlife Refuge 2016). We know of no published studies, however, that have compared this method with frequent checks of individual nests.

Variability in the timing of breeding relative to survey dates will affect reproductive success estimated from 2

surveys (Southwell and Emmerson 2015). An understanding of the phenology of the species or colony is required to choose appropriate dates on which to count nests and chicks or to correct counts done on suboptimal dates (Emmerson et al. 2003, Lynch et al. 2009, Southwell et al. 2010).

Alternatively, a colony may be visited or nests checked at frequencies less than daily throughout a season, for logistical reasons or because the species is sensitive to disturbance (Carey 2009). If all eggs in a nest are laid and lost in the interval between nest checks, the nest will not be detected as a nest with eggs and will unintentionally be excluded from the calculation of reproductive success. This exclusion artificially inflates estimates of reproductive success because 100% of these nests failed (Mayfield 1975).

We used 33 yr of data for the Magellanic Penguin (*Spheniscus magellanicus*), a long-lived colonial seabird, to determine the effects of sampling frequency on estimates of reproductive success. Our objectives were to compare data on reproductive success collected at a fine temporal scale (daily nests checks) with data collected at coarser scales (2 surveys per season or less frequent nest checks) and to explore potential sources of bias in reproductive success estimated from 2 surveys per season (Table 1). We hypothesized that (1) reproductive success estimated from 2 surveys would be sufficiently accurate for population models; (2) counts from 2 surveys could be adjusted based on data easily collected on the survey days; (3) temporal trends in reproductive success over multiple breeding seasons would be detectable from 2 surveys per season; and (4) less-frequent nest checks during laying and incubation would overestimate reproductive success compared with daily nest checks.

METHODS

Study Species

Magellanic Penguins breed in colonies in Argentina, Chile, and on the Falkland (Malvinas) Islands (Schiavini et al. 2005, Boersma et al. 2013). They are migratory (Boersma et al. 1990, García-Borboroglu et al. 2006, Stokes et al. 2014), arriving at colonies in Argentina in September or October. Males arrive before females and secure nest sites. Females lay 2 eggs, usually in October, and usually take the first long incubation shift while males go to sea to forage after 1 or both eggs are laid (Boersma et al. 1990). Magellanic

TABLE 1. Sources of bias in seabird reproductive success estimated from 2 surveys per season, and the consequences for occupied-nest counts, chick counts, and estimates of reproductive success.

Potential source of bias	Occupied nests in October	Chicks in January	Reproductive success
Breeders not detected: Arrived late or left early Temporarily absent Present but not seen	Undercount		Overestimate
Occupied nests in which eggs are not laid	Overcount		Underestimate
Counting error	Undercount or overcount	Undercount or overcount	*
Breeding synchrony low	Undercount	Undercount	*
Spatial heterogeneity in reproductive success		Undercount or overcount	Underestimate or overestimate
Chicks fledge before survey		Undercount	Underestimate
Chicks die after survey		Overcount	Overestimate
Chicks move out of plot		Undercount	Underestimate
Chicks move into plot		Overcount	Overestimate

* The effect on reproductive success depends on the relative amount of undercounting and/or overcounting for nests and chicks.

Penguins lay 1 clutch per year and the laying of replacement clutches is rare if eggs are lost (Boersma and Rebstock 2014). When males return from their long foraging trip, they take a long incubation shift while females forage. Mates then alternate short foraging trips with incubating eggs or brooding chicks. Chicks hatch in November or early December (Boersma et al. 1990) and parents take turns guarding the chicks for about a month before leaving the chicks unguarded while both parents forage. Unattended chicks may remain in their nests or move to a nearby nest with or without other chicks. Chicks fledge in January or February (Boersma et al. 2013). Starvation is the most common cause of chick mortality in most years ($39 \pm 18\%$), followed by predation ($9 \pm 4\%$). In some years, rainstorms kill more chicks (up to 50%) than any other cause of mortality (Boersma and Rebstock 2014).

Field Methods

We followed Magellanic Penguin nests at Punta Tombo, Argentina (44.045°S , 65.223°W), to determine reproductive success each season from 1983 (the 1983–1984 season) through 2015, except in 2011 (32 seasons). The colony occupied approximately 400 ha (Rebstock et al. 2016). In ~ 80 ha of the colony (Figure 1), we checked nests occupied by banded penguins every day or every other day during egg laying and hatching, and every 6–10 days the rest of the season. In a study area of ~ 0.7 ha within the larger study area (Figure 1), we checked nest sites daily prior to laying and as long as they had eggs or chicks, and banded all penguins. In both cases we started checking nests in late September or early October, before egg laying, except in 2010 (checks started in late October), and continued to check all study nests until they lost their eggs or chicks, or their chicks fledged. We marked chicks to

identify nest number and hatching order. We counted a chick as having fledged if it was not found dead and weighed at least 1,800 g after January 9 (Boersma et al. 1990). We calculated reproductive success each year as the number of chicks that fledged divided by the number of nests with eggs (maximum eggs per nest = 2).

We surveyed 51 plots at Punta Tombo (Figure 1) twice each season (in October and January) from 1994 to 2013, except in 1995 and 2011 (18 seasons; Rebstock et al. 2016). Plots were 100 m², permanently marked with rebar stakes, and spaced every 100 m over several kilometers of the colony. In October surveys, we counted nests with penguins or eggs as occupied nests. We calculated reproductive success each season as the total number of chicks in the January survey divided by the total number of occupied nests in the October survey. We used occupied nests rather than nests with eggs because there were few eggs when we conducted October surveys in most years. This was a major source of bias (see below).

We estimated peak fledging dates each season from 1989 to 2015, except in 2011 (26 seasons), by counting fledglings leaving the colony at 1 landing beach for 18 to 110 min starting before 0830 hours. We counted every day or every other day, starting between January 1 and 17, and continuing until February 11 to March 1. The peak fledging date was the date with the highest number of fledglings leaving the colony per hour.

Accuracy of Reproductive Success Estimated from 2 Surveys

To be useful in demographic models, parameter estimates must be within a target percentage of the true value, depending on the sensitivity of the model to the parameter (Beissinger and Westphal 1998). We calculated the percent

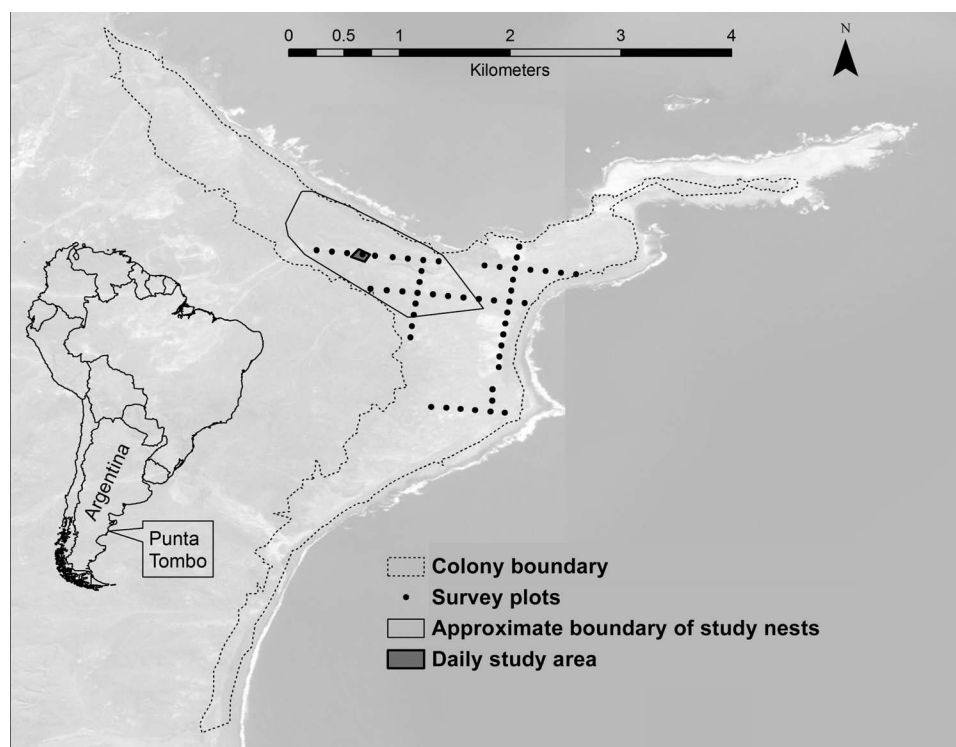


FIGURE 1. Map of the Magellanic Penguin colony at Punta Tombo, Argentina. The colony occupies ~ 400 ha, the study nests ~ 80 ha, and the daily study area ~ 0.7 ha. Study nests (including nests in the daily study area) were followed in each season that they were occupied by study birds from 1983 through 2015. The 51 survey plots were 100 m^2 (5.64-m radius circles) spaced 100 m apart, and were surveyed twice each breeding season from 1994 through 2013. The inset of South America shows the location of Punta Tombo.

difference between reproductive success (RS) estimated from nest checks and surveys each season as $100 \times (|RS_{\text{SURVEYS}} - RS_{\text{NEST CHECKS}}|) / RS_{\text{NEST CHECKS}}$, where $||$ indicates the absolute value of the difference. We set the target percent difference in reproductive success as $<15\%$, based on the changes in reproductive success that were calculated to change population growth rates by 1% for other seabird species (Appendix Table 4). We chose 15% as the target, rather than the lowest value that we found in the literature (11%), to avoid being too restrictive for most species. If a species' population growth rate is more sensitive to reproductive success, extra attention should be paid to conducting counts on optimal dates and accounting for egg and chick losses (see below).

We compared reproductive success estimated for each year from the surveys with that from nest checks using linear regression. We used nest checks in the 80-ha study area ($n = 388 \pm 86$ (SD) nests with eggs per year) because the spatial scale matched the scale of the surveys better than the scale of the 0.7-ha daily study area (Figure 1). We also compared long-term mean reproductive success estimated from surveys and nest checks using a t -test paired by year (18 samples over 20 yr).

Nests with eggs was a better variable than occupied nests for calculating reproductive success when nests

were counted on optimal dates. We determined the best dates on which to count chicks and nests with eggs using the daily study area ($n = 129 \pm 70$ (SD) nests per year) and counts from 1983 to 2015. We counted the number of males, females, chicks, and nests with eggs each day in this study area and counted the total number of nests in which eggs were laid and the number of chicks that fledged. For each season, we determined the maximum count of nests with eggs over each 5-day period from October 1 to November 29 and the maximum count of chicks each 5-day period from December 27 to February 19. We divided the maximum count of nests with eggs for each period by the total number of nests with eggs. We divided the maximum chick count for each period by the actual number of chicks fledged. The best periods to count nests with eggs and to count chicks were the periods with the ratios closest to one and the lowest coefficients of variation (CV), indicating that the period was consistently representative. We calculated reproductive success using the maximum counts from the optimal periods each season. No counts occurred on the optimal dates in 1990 and 1991. We compared this measure of reproductive success with that calculated from the daily nest checks.

Adjusting Counts

In some occupied nests counted in October surveys, eggs were not laid because the males did not get mates (Table 1). We hypothesized that the percentage of occupied nests in which eggs were laid would be higher in years of high reproductive success than in years of poor success. We regressed the proportion of occupied nests in which eggs were laid on reproductive success from nest checks. Occupied nests were counted in the daily study area on the October survey dates. We included reproductive success squared to test for a quadratic relationship. We hypothesized that the percentage of occupied nests in which eggs were not laid would be higher for earlier surveys, and that survey date could be used to adjust reproductive success estimated from 2 surveys. We regressed the percentage of occupied nests in which eggs were not laid on the survey date (date in October). We included survey date squared to test for a quadratic relationship.

Southwell et al. (2010) developed an 'availability adjustment factor' to correct counts made before or after the optimal dates for Adélie Penguins (*Pygoscelis adeliae*). We used a modified version of this factor, using 32 yr of data. We divided the count of nests with eggs in the daily study area on each day in October and November by the maximum count for the respective year. We calculated an adjustment factor for each day from October 1 to November 30 as the mean across years for that date (61 adjustment factors). We divided each count of nests with eggs by the adjustment factor for the corresponding day, for all days with nonzero counts. If the method works, each adjusted count should be close to the maximum count for that year.

We overestimated the number of chicks that fledged because some of the chicks counted did not fledge, but died after the survey (Table 1). We hypothesized that the percentage of chicks that died after the January survey date would be higher in years of low reproductive success. We regressed the proportion of chicks in the daily study area that died after the January survey dates on reproductive success from nest checks and included reproductive success squared to test for a quadratic relationship.

We adjusted the estimate of reproductive success calculated from 2 counts in the daily study area on the optimal dates (nests with eggs in October, and chicks in January) by removing the percentage of chicks that were unlikely to fledge. We weighed a sample of chicks between January 26 and 31 each season ($n = 237 \pm 138$ (SD) chicks) and calculated the proportion of chicks that weighed at least 1.5, 1.6, 1.7, 1.8, 1.9, and 2.0 kg. We multiplied reproductive success estimated from the 2 counts in the daily study area by each proportion and compared the adjusted estimates with reproductive success calculated from the daily nest checks.

Other Potential Sources of Bias

Chick movements. Chicks sometimes move to other nests or wander outside nests, potentially causing detection bias for chicks in the January survey (Table 1). We determined the maximum distance that chicks moved from their nests (prior to fledging) and the timing of their movements. We used 4,497 chicks from the daily study area in 1983–2015. If chicks frequently moved more than a few meters, then they were likely to move in and out of our survey plots (11.28-m-diameter circles).

Counting error. As a test of detection bias and counting error in January surveys, we surveyed 24 permanently marked plots at Cabo dos Bahías (44.913°S, 65.547°W), a colony ~100 km south of Punta Tombo, on January 5 and again on January 16, 2016, and compared chick counts between the 2 dates. Likely no or few chicks fledged at Cabo dos Bahías between January 5 and 16. We saw the first chicks fledging at Punta Tombo on January 19, and chicks at Cabo dos Bahías and Punta Tombo were generally at about the same developmental stage in mid-January, with juvenile plumage not fully grown.

Habitat heterogeneity. Reproductive success varies with habitat type at Punta Tombo (Stokes and Boersma 1998, Rebstock et al. 2016) and in other seabird colonies (Hamer et al. 2002). If the habitats sampled in the survey are not representative of the colony as a whole, reproductive success will be overestimated or underestimated. We counted the number of each type of nest used by Magellanic Penguins (burrows, bushes, and scrapes; Stokes and Boersma 1991) in our surveys and in nest checks and compared the frequencies of nest types in surveys and nest checks with a χ^2 test. To determine whether differences in the percentages of each nest type between surveys and nest checks affected reproductive success, we calculated the mean reproductive success for 5,011 burrow nests, 8,004 bush nests, and 40 scrape nests for all years. We then multiplied the mean reproductive success for each nest type by the percentage of that nest type in nest checks and in surveys, and compared the sums between nest checks and surveys.

Trend Detection

We checked for trends in reproductive success over time by regressing reproductive success on year and year squared, using both surveys and nest checks. Reproductive success did not increase or decrease over time using either method (see Results). Because negative results do not make for a satisfying comparison of methods, we then randomized the order of years without replacement 1,000 times, using the same random order for both methods in each repetition, and ran the regressions for each repetition. We counted the number of repetitions that resulted in each possible outcome (agreement or disagreement between nest checks and surveys). We also tested whether

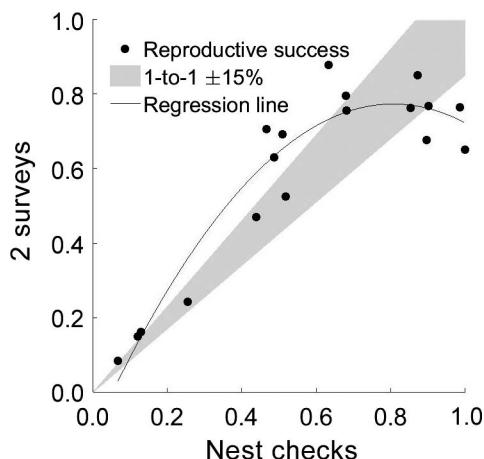


FIGURE 2. Reproductive success of Magellanic Penguins at Punta Tombo, Argentina, 1994–2013, estimated from 2 surveys per season, 1 in October to count occupied nests and 1 in January to count chicks, was nonlinearly correlated with reproductive success calculated from nest checks. The quadratic regression line is shown ($R^2 = 0.90$). The shaded area is the 1-to-1 relationship $\pm 15\%$. Points outside the shaded area indicate years when the 2 methods of estimating reproductive success differed by $\geq 15\%$.

the slopes were similar (whether the slope of reproductive success calculated from nest checks was within the 95% confidence interval of the slope estimated from surveys).

Nest-check Frequency

We investigated how reproductive success calculated from nest checks would be affected by less-than-daily nest checks during egg laying using 10 yr (1998–2005 and 2007–2008) of data from our daily study area when no counts were missed between October 1 and December 15. We selected nest checks conducted every 2 days to every 30 days during laying and incubation, removed the counts between the selected days, and calculated reproductive success using the maximum count of nests with eggs on the selected checks.

Statistics

We used Stata 11.2 (StataCorp, College Station, Texas, USA) for statistical tests. We wrote custom programs in MATLAB R2013b (The MathWorks, Natick, Massachusetts, USA) for randomization procedures. We present means \pm SE unless otherwise stated.

RESULTS

Accuracy of Reproductive Success Estimated from 2 Surveys

Two surveys usually underestimated reproductive success in good years and overestimated it in average years compared with nest checks (Figure 2). The correlation

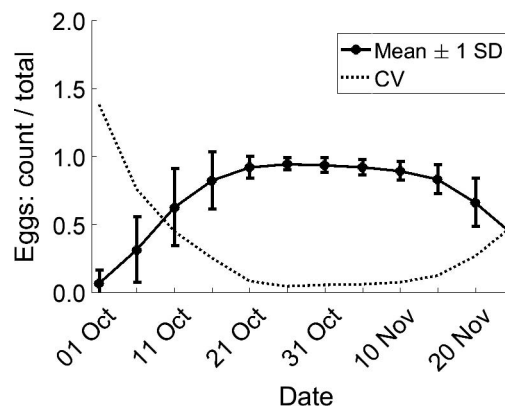


FIGURE 3. Maximum count of nests with eggs of Magellanic Penguins at Punta Tombo, Argentina, for each 5-day period from October 1 through November 29 divided by the total number of nests with eggs each season, 1983–2015, for nests checked daily. The best period for counting nests with eggs was October 26–30, with the mean proportion closest to 1 and the lowest coefficient of variation (CV). Each tick marks the start of a 5-day period.

was significant ($F_{1,16} = 40.9$, $P < 0.001$, $R^2 = 0.70$), but the fit was better when the square of reproductive success from nest checks was included ($F_{2,15} = 74.2$, $P < 0.001$, $R^2 = 0.90$), confirming that the relationship was nonlinear.

The long-term means matched well, but the agreement between the 2 methods was poor in most years. Long-term mean reproductive success calculated from nest checks was 0.58 ± 0.07 , compared with 0.59 ± 0.06 calculated from surveys ($t_{17} = 0.1$, $P = 0.92$). The annual difference was $<15\%$ in only 7 of 18 yr (Figure 2). The mean difference was 21%.

The optimal 5-day period for counting nests with eggs was October 26–30, and dates from October 21 through November 14 were also good (Figure 3). The optimal 5-day period for counting chicks was January 26–30 (Figure 4). The ratio of chicks counted to actual chicks fledged was also close to 1 for January 21 through February 4, but CVs increased rapidly during February as more chicks fledged.

When we used counts done on the optimal dates in late October (nests with eggs) and January (chicks) at Punta Tombo, the correlation and the mean percent difference between reproductive success estimated from 2 counts and nest checks improved. The correlation was linear ($F_{1,28} = 738$, $P < 0.001$, $R^2 = 0.96$). The squared term in the quadratic regression was not significant ($t = 1.0$, $P = 0.32$). The difference between the 2 methods was $<15\%$ in 19 of 29 yr (reproductive success from nest checks was 0 in 1984, so the percent difference was undefined), with a mean of 11%. The long-term mean reproductive success calculated from counts on the optimal dates (0.53 ± 0.05), however, was significantly higher than the mean repro-

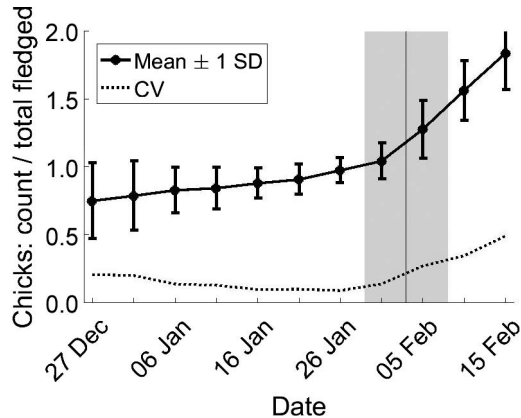


FIGURE 4. Maximum count of chicks of Magellanic Penguins at Punta Tombo, Argentina, for each 5-day period from December 27 through February 19 divided by the total number of chicks that fledged each season, 1983–2015, for nests checked daily. The best period for counting chicks was January 26–30, with the mean proportion closest to 1 and the lowest coefficient of variation (CV). Each tick marks the start of a 5-day period. The vertical line and shading mark the median date (February 3) and median absolute deviation of peak fledging dates from 1989 to 2014.

ductive success estimated from daily nest checks (0.48 ± 0.05 ; $t_{29} = 4.5$, $P < 0.001$).

Adjusting Counts

The percentage of occupied nests on the October survey dates into which eggs were laid in that season ranged from 52% to 114%. In most years with high reproductive success, the percentage of occupied nests in which eggs were laid was low (Figure 5). The relationship between the proportion of nests in which eggs were laid and reproductive success was not quite significant ($F_{1,16} = 4.2$, $P = 0.06$). If we excluded 1996, an outlier in the relationship, from the analysis, the relationship was negative and linear ($F_{1,15} = 9.5$, $P = 0.008$, $R^2 = 0.35$).

Survey date could not be used to adjust counts of occupied nests. The percentage of nests in which eggs were not laid was not related to the date of the counts ($F_{1,16} = 0.05$, $P = 0.83$). The squared term was also not significant ($t = 0.01$, $P = 0.99$). Few eggs were laid in the first week of October, when we conducted surveys in most years.

The adjustment factors for each day for the counts of nests with eggs based on the mean percentage of the maximum counts were not helpful outside the period when the number of nests with eggs was close to the total for that year (and adjustment factors were least needed). The means of the adjusted counts were $<10\%$ different from the total nests with eggs, as expected, for all days except October 1–7, when few eggs were laid. Individual adjusted counts, however, matched maximum counts only

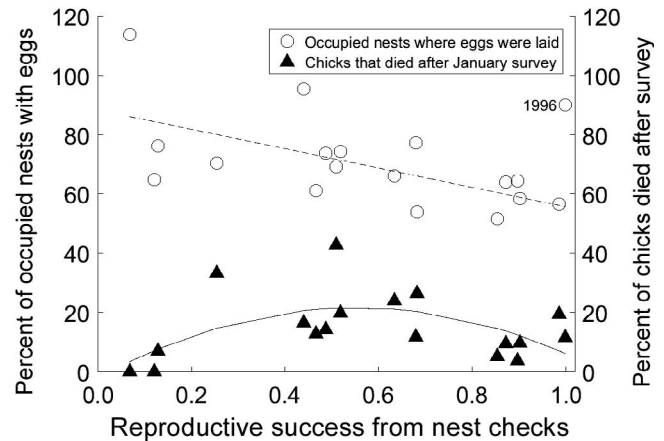


FIGURE 5. The percentage of occupied nests in which eggs were laid in a daily study area during the October survey dates was linearly related to reproductive success calculated from nest checks (dashed regression line) in 1983–2015 if 1996 (labeled point) was excluded. In 2000, the percentage was >100 because Magellanic Penguins returned to the colony at Punta Tombo, Argentina, very late. Many nests were empty on the survey date, but later had pairs and eggs. The percentage of chicks in the daily study area that survived until the January survey date but later died was nonlinearly related to reproductive success calculated from nest checks (solid regression line).

a little better than the unadjusted counts. For the period of October 28 to November 9, 99% of adjusted counts were within 15% of the maximum counts, compared with 96% of unadjusted counts. Outside that period, 44% of adjusted counts were within 15% of the maximum counts, compared with 35% of unadjusted counts.

The percentage of chicks in our daily study area that survived until the January survey date but later died ranged from 0% in 2000 and 2012 to 43% in 2010. The percentage of chicks that died after the survey was nonlinearly related to reproductive success across years (Figure 5; $F_{2,15} = 3.5$, $P = 0.05$, $R^2 = 0.23$, squared term: $t = 2.7$, $P = 0.02$). A smaller percentage of chicks died late in bad and good years than in average years. In the worst years, 2000 and 2012, almost all chicks died in early to mid-December, leaving fewer than 5 chicks in the study area in January, all of which fledged.

When we adjusted reproductive success estimated from 2 counts (nests with eggs, and chicks) conducted on optimal dates by the proportion of chicks that weighed at least 1.5–1.7 kg, the long-term means were similar to mean reproductive success estimated from nest checks (Table 2). For weights of 1.8–2.0 kg, the estimated long-term means of reproductive success were significantly lower than the mean reproductive success calculated from nest checks (Table 2). The best weight to use for removing chicks unlikely to fledge from the counts was 1.7 kg; using this weight resulted in the most years with differences $<15\%$ and a mean difference of 9% (Table 2). The regression of

TABLE 2. We adjusted counts of Magellanic Penguin chicks for reproductive success estimates from 2 surveys at Punta Tombo, Argentina, 1983–2015, by subtracting the percentage of chicks that weighed less than 1.5, 1.6, 1.7, 1.8, 1.9, and 2.0 kg. When we used weights of 1.5–1.7 kg, the long-term mean estimate of reproductive success from 2 surveys was similar to the mean calculated from nest checks (0.48 ± 0.05 ; *t*-test statistics). Using 1.7 kg resulted in the most years with a difference in reproductive success of <15% between the 2 methods.

Minimum chick mass (kg)	Long-term mean \pm SE (<i>t</i> -test statistics, <i>df</i> = 29)	Years with difference <15% (maximum = 29)	Regression statistics (<i>df</i> = 1, 28)
1.5	0.50 ± 0.05 (<i>t</i> = 1.7, <i>P</i> = 0.10)	23	<i>F</i> = 626, <i>P</i> < 0.001, <i>R</i> ² = 0.96
1.6	0.49 ± 0.05 (<i>t</i> = 0.4, <i>P</i> = 0.66)	24	<i>F</i> = 550, <i>P</i> < 0.001, <i>R</i> ² = 0.95
1.7	0.48 ± 0.05 (<i>t</i> = 0.7, <i>P</i> = 0.50)	25	<i>F</i> = 464, <i>P</i> < 0.001, <i>R</i> ² = 0.94
1.8	0.46 ± 0.04 (<i>t</i> = 2.1, <i>P</i> = 0.05)	21	<i>F</i> = 387, <i>P</i> < 0.001, <i>R</i> ² = 0.93
1.9	0.44 ± 0.04 (<i>t</i> = 3.2, <i>P</i> = 0.003)	20	<i>F</i> = 277, <i>P</i> < 0.001, <i>R</i> ² = 0.90
2.0	0.44 ± 0.04 (<i>t</i> = 4.2, <i>P</i> < 0.001)	19	<i>F</i> = 202, <i>P</i> < 0.001, <i>R</i> ² = 0.87

adjusted reproductive success from 2 counts on reproductive success from nest checks was significant and linear (Figure 6, Table 2).

Other Potential Sources of Bias

Chick movements. Some chicks wandered from their nests before they were ready to fledge, but they usually did not wander far until late in the season (February). We found 23% of chicks out of their nests at least once (1,036 of 4,497 chicks). The mean distance moved was ~ 3 m for the 287 chicks for which we estimated the distance moved. Most (98%) moved ≤ 10 m, and 58% moved ≤ 2 m. Hence, $\sim 10\%$ of all chicks moved > 2 m from their nests. Chicks as young as a few days old moved (beginning November 13), especially if they were abandoned or were in a nest that flooded. The mean distance moved increased from December to March (Table 3). The maximum distance moved was 30 m. Of the 42 nest checks in which chicks were found > 10 m from their nests, 4 of the chicks moved between January 24 and 30 and the rest moved in February.

Counting error. Detection bias and counting error in January surveys were minimal. Chick counts were similar at Cabo dos Bahías in the January 5 and January 16 surveys ($F_{1,22} = 271$, *P* < 0.001, *R*² = 0.92, slope = 1.07 ± 0.06). The mean of the first count was 4.3 ± 0.9 chicks per plot, and the mean of the second count was 4.5 ± 1.0 chicks per plot.

Habitat heterogeneity. Burrow nests (46%) and scrapes (0.8%) were overrepresented in surveys compared with nest checks (44% and 0.2%, respectively; $\chi^2 = 30.5$, *P* < 0.001). When scrapes were excluded because of the small sample size, burrow nests were still significantly overrepresented in surveys ($\chi^2 = 4.0$, *P* = 0.05). The small differences in the proportions of nest types detected in surveys vs. nest checks did not bias the mean calculated reproductive success estimate, however. Mean reproductive success was 0.56 for burrow nests, 0.47 for bush nests, and 0.30 for scrape nests. Multiplying the mean reproductive success for each nest type by the proportion of

nests of that type and summing gave 0.509 for nest checks and 0.510 for surveys (0.2% difference).

Trend Detection

Two surveys per season were sufficient to detect long-term trends. There was no linear trend in estimated reproductive success over time using nest checks ($F_{1,16} < 0.01$, *P* = 0.99) or surveys ($F_{1,16} = 0.4$, *P* = 0.55), and the squared term was not significant for nest checks (*t* = 1.0, *P* = 0.35) or surveys (*t* = 0.9, *P* = 0.41). Had there been a long-term trend, however, we would likely have detected it using 2 surveys. The presence or absence of a significant trend from 2 surveys matched the results from nest checks in 95% of 1,000 randomizations. Both methods detected

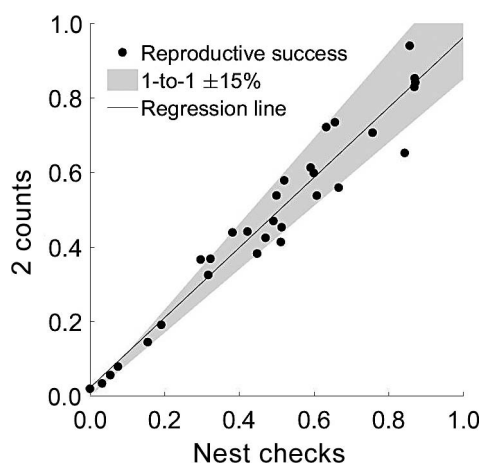


FIGURE 6. Reproductive success of Magellanic Penguins at Punta Tombo, Argentina, 1983–2015, estimated from 2 counts per season, 1 in late October (nests with eggs) and 1 in late January (chicks), was linearly correlated with reproductive success calculated from daily nest checks in the same study area. Reproductive success from the 2 counts was multiplied by the proportion of chicks that weighed ≥ 1.7 kg on January 26–31, weighed anywhere in the colony. The linear regression line is shown (*R*² = 0.94). The shaded area is the 1-to-1 relationship $\pm 15\%$. Points outside the shaded area indicate years when the 2 methods of estimating reproductive success differed by $\geq 15\%$.

TABLE 3. The mean distance that Magellanic Penguin chicks moved from their nests increased from December to March, 1983 to 2015, at Punta Tombo, Argentina ($F_{3,1673} = 39.5$, $P < 0.001$). N total is the number of nest checks for which a chick moved, whether or not we estimated the distance. N with distance is the number of nest checks for which we estimated the distance that a chick moved. The number of chicks that moved increased in late December, after chicks were left alone while both parents foraged. Numbers decreased in February and March as chicks fledged or died.

Month	N total	N with distance	Mean distance \pm SE (m)
November	40	0	N/A
December	1,492	118	1.2 ± 0.1
January	4,682	802	2.5 ± 0.1
February	2,313	736	3.6 ± 0.1
March	32	21	4.2 ± 0.5

trends in 24 randomizations, and neither method detected a trend in 930 randomizations. In all 24 randomizations in which a trend was found by both nest checks and surveys, the trends were in the same direction (11 positive and 13 negative) and had similar slopes. In all cases, whether both methods or only one method showed a significant trend, the slope of reproductive success calculated from nest checks was within the 95% confidence interval for the slope calculated from surveys.

Nest-check Frequency

The maximum number of nests with eggs and reproductive success were the same whether nests were checked every day or every other day in our test of nest-check frequency. Fewer nests with eggs were counted if nests were checked less frequently than every 2 days, overestimating reproductive success. Checking nests every 3 days to every 8 days increased reproductive success estimates by $<5\%$, and checking nests every 9 days to every 19 days increased reproductive success estimates by $<15\%$. Only in 2000, a year with delayed breeding and very low reproductive success, did differences exceed 15% for checks conducted every 20–25 days vs. daily. The estimate of reproductive success did not increase steadily as nest-check frequency decreased, but increased then decreased depending on the timing of checks relative to egg laying and loss (Figure 7).

DISCUSSION

Collecting detailed data on egg laying and nest fate throughout one or more breeding seasons is time consuming and costly, and some species are sensitive to the disturbance caused by checking nests (Carey 2009). Time-lapse photography and unmanned aerial vehicles (drones) are increasingly used to estimate population size

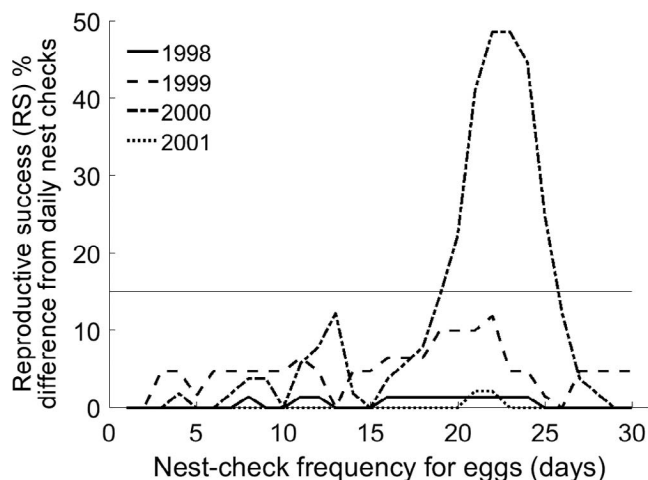


FIGURE 7. Estimates of reproductive success for Magellanic Penguins at Punta Tombo, Argentina changed irregularly as the frequency of nest checks decreased from daily to every 30 days during egg laying and incubation. We show the percent difference between reproductive success for each nest-check frequency from 2 to 30 days and reproductive success from daily checks. The horizontal line is 15%, our target level of accuracy. We selected 4 yr for display, including the only year with a difference $>15\%$ (2000) and the years with the lowest differences (1998 and 2001).

and reproductive success (Southwell and Emmerson 2015, Hodgson et al. 2016, Merkel et al. 2016), but do not work for species that nest in burrows or crevices, under vegetation, or in low-density, dispersed colonies. We evaluated the use of 2 surveys per breeding season as a quicker, less expensive, and less invasive alternative to following nests throughout a season for estimating reproductive success. We found that 2 surveys done on optimal dates during incubation and chick rearing, with an adjustment for the proportion of chicks likely to fledge, yielded estimates of reproductive success that were sufficiently accurate for population models in most years. These estimates were also reliable for long-term means and trend detection.

Accuracy of Reproductive Success Estimated from 2 Surveys

Survey date and the necessity of using occupied nests instead of nests with eggs were the main sources of bias in our data, as we conducted October surveys before peak counts of nests with eggs. In October we counted nests with single males as occupied, and some of those males did not get mates that season. Females return to the colony later than males each season (Boersma et al. 1990), and many males never get mates because there are more males than females at Punta Tombo (Boersma et al. 2013). If the sex ratio of a colony is skewed, or there are nonbreeders present (e.g., Waugh et al. 2015), conducting counts of nests with eggs, rather than occupied nests, on optimal

dates will improve estimates of reproductive success. Selecting the best dates to count, based on the phenology of the study species, is a critical consideration when designing surveys to estimate reproductive success.

Chicks that died after the January survey were the other large source of bias in most years. The timing of chick fledging and especially chick death is likely a greater source of bias than the timing of egg laying, even if the first survey is conducted after peak egg laying. In Adélie Penguins, counts of chicks were much more variable than counts of adults and occupied nests during incubation, and reproductive success and fledgling mass were more variable than adult weights (Emmerson et al. 2003). A good count of nests with eggs could be made on one day, were egg predation to be low. Except in some years with early and nearly complete nest failure, it was not possible to get an accurate count of fledglings on any one day in our study. A count just prior to peak fledging, combined with information on chick mass that indicated the percentage of chicks that was likely to fledge, reduced this source of bias.

Peak dates for egg and chick numbers vary from year to year in most seabirds (Birkhead and Nettleship 1980, Emmerson et al. 2003, Barbraud and Weimerskirch 2006, Lewis et al. 2012), so counting on fixed dates will introduce error into calculations in some years, as this study showed. In Magellanic Penguins, the date of the maximum number of nests with eggs varied by almost 3 weeks (P. D. Boersma personal observation). In a few years, counting on optimal dates would miss the peak of nests with eggs. Modeling dates of peak egg numbers using environmental covariates may help to determine the best dates on which to count (Lynch et al. 2009). Optimal dates could then be determined in advance each season, if there were a strong correlation between phenology and pre-season environmental data that are available remotely (e.g., weather-station data, satellite sea-surface temperature).

Even without environmental data, it may be possible to tell if breeding in a given year started very early or very late during a short visit to the colony, because the short visit will likely be after the start of the breeding season. Laying dates can be estimated from the density of eggs because eggs lose water during incubation (Furness and Furness 1981, Demongin et al. 2007). The size of chicks (Demongin et al. 2007) or the ratio of incubating males to females (Southwell et al. 2010) may also indicate if eggs were laid earlier or later than usual. If the breeding season appears to be exceptionally early or late, the data should not be used in demographic models, as they are likely not representative of mean reproductive success.

We used simple linear regression rather than a mixed-effects model, even though some individuals and pairs bred in more than 1 yr. This would have affected the *P*-value of the regression, but not the R^2 value or slope, and the

regression would have been significant with as few as 1 or 2 degrees of freedom. We did not count all the same pairs every season as up to 36% of females and 23% of males skipped breeding in some years (Boersma and Rebstock 2010) and 30% of females and 21% of males changed nests between 2 consecutive seasons (Boersma et al. 2013). Nest type (Stokes and Boersma 1998, Rebstock et al. 2016), weather (Boersma and Rebstock 2014), and prey availability (Boersma and Rebstock 2009) strongly influence reproductive success, regardless of breeder quality. Studies of seabird colonies without marked birds cannot account for individual ID when comparing reproductive success among years.

Adjusting Counts

Adjustment factors are often recommended when breeders or chicks are counted on suboptimal dates (Lynch et al. 2009, Southwell et al. 2010, CCAMLR 2014). In our study, adjusting counts of nests with eggs based on the mean (over 32 yr) of the percent of the maximum count for the year did not improve estimates in most cases. We did not use a smoothing function to create the adjustment factors (Southwell et al. 2010) because the long-term means were sufficiently smooth. Variability in egg laying dates caused the error in counts of nests with eggs, and additional information is needed to adjust counts done on suboptimal dates. Nests with eggs should be counted instead of occupied nests when some adults present do not get mates, or when eggs are not laid in all the occupied nests.

Chick counts required an adjustment even if done on optimal dates because chicks continued to die after the survey. The timing of chick mortality depends on food availability and extreme weather (Boersma and Stokes 1995, Boersma and Rebstock 2009, 2014). Older chicks (more than ~50 days of age) were less likely to be killed by predators or die in rainstorms, but chicks continued to die from starvation and heat up to 70–100 days of age, throughout January and February (Boersma and Rebstock 2014). For Magellanic Penguins, chick mass in late January was a good basis for adjusting chick counts. For other species, other factors, such as body size or juvenile plumage, may be more appropriate, or no adjustment may be needed if few chicks die late in the season.

Minor Sources of Bias

Chick movements. Chicks moving into or out of survey plots was a small source of error as most chicks (77%) did not move and those that did generally moved short distances. Chicks from nests close to the edge of a plot may move out of or into the plot, but if the habitat is the same, the movements should be random. This source of bias can be minimized by not surveying plots at the boundary between 2 habitat types where nest quality varies.

Counting error. Counting error was small, as our repeated surveys in October (Rebstock et al. 2016) and January (this study) showed. The same observers usually did the October and January surveys within a season, reducing any observer bias in calculating reproductive success.

Detection bias can be caused by breeders or chicks not being present or not being seen if present (Southwell et al. 2010). Seabirds that nest in burrows or crevices can be particularly difficult to detect (Reynolds and Renner 2014). Failure to detect breeders was at most a minor source of bias in our data. In Magellanic Penguins, once a male acquires a nest he defends it until eggs are laid if he gets a mate, or until well into the egg-laying period if he does not (Boersma et al. 1990, Renison et al. 2002, Boersma et al. 2013). A few pairs lose their eggs early in the season, but failed breeders frequently stay in their nests for days to weeks if eggs are lost early (P. D. Boersma personal observation). Thus, occupied nests are generally continuously occupied.

Reproductive success of Adélie Penguins calculated from automatic-camera images was underestimated relative to nest checks when early-January images were used because small chicks that were still being brooded were not seen (Southwell and Emmerson 2015). We looked in every nest in each plot and lifted penguins to see if they were incubating eggs or brooding chicks. Most nests were easy to check, so low detection was not a problem in our egg or chick counts.

A few chicks fledged before our second survey, which was always done before peak fledging each season. However, we do not recommend conducting the second survey before chicks begin fledging because this could increase the bias due to chicks dying after the survey. We showed that late January, close to peak fledging, was the best time to count chicks at Punta Tombo in most years.

Habitat heterogeneity. There was only partial overlap between our survey plots and nests that we checked. The nests that we checked changed somewhat from year to year because many study nests depended on where known-age penguins chose to breed. The small differences in the percentage of each nest type in surveys vs. nest checks did not affect reproductive success. This source of bias can be minimized by ensuring that survey plots (and nests checked) represent the breeding-habitat quality of the colony as a whole.

Trend Detection

For detecting trends, an index or proxy of the variable suffices (Greenwood and Robinson 2006). We showed that long-term trends in reproductive success could be detected with 95% accuracy, even with biased estimates from suboptimal dates. The slopes were always in the same

direction and always similar, even though the 2 methods disagreed in a few of the randomizations at $\alpha = 0.05$.

Nest-check Frequency

We showed that the overestimation of reproductive success was generally small (<15%) for nest checks conducted every 3–30 days. In general, the optimal nest-check frequency depends on the length of the peak egg laying period and the timing of counts relative to the peak. Less frequent checks could produce a more accurate estimate of reproductive success than more frequent checks if a less frequent check were to occur at the time of peak numbers of nests with eggs. Counting several times during a season can give more certainty about when the peak of nests with eggs occurred in a given year. Egg-loss rates should also be taken into account. Higher egg-loss rates require more frequent nest checks to ensure that all nests with eggs are found. When nests are checked infrequently or nests are found after eggs have been laid, the Mayfield method (Mayfield 1975) or logistic-exposure model (Shaffer 2004) can be used to calculate reproductive success.

Recommendations

Our general results are applicable to other colonial species that breed synchronously and lay a single clutch per year and do not lay a replacement clutch if the first clutch is lost. We found that 2 surveys on optimal dates, with an adjustment for the percentage of chicks that are likely to fledge, can produce accurate estimates of reproductive success that can be used in population models and trend detection. The timing of chick mortality will affect the need for an adjustment factor; if few chicks die late in the season, an adjustment will not be necessary. Counts of nests with eggs will be difficult to adjust because of variability in laying dates. Knowledge of the phenology and egg- and chick-loss rates of a species will be necessary to estimate reproductive success accurately from two surveys. Variability in egg laying dates among years will result in occasional inaccurate estimates when counting on fixed dates, but additional data or observations taken at the time of the surveys can indicate whether this may be a problem in a given year.

ACKNOWLEDGMENTS

The Penguin Project is affiliated with the Wildlife Conservation Society (WCS), the Global Penguin Society (GPS), the University of Washington, and the National Scientific and Technical Research Council of Argentina (CONICET). We thank the Province of Chubut and the La Regina family for access to the penguin colony. We thank the La Regina family, Carlos García, Popi García-Borboroglu, Graham Harris, Patricia Harris, Carol Passera, and Centro Nacional Pa-

tagónico, Argentina (CENPAT), for logistical support, and the many students and field volunteers who collected data over the years, including volunteers through the Wildlife Conservation Society and students from Puerto Madryn and Trelew, Argentina. Comments from 2 anonymous reviewers improved the manuscript.

Funding statement: Funds for research were provided by the WCS, Exxonmobil Foundation, the Pew Fellows Program in Marine Conservation, the Disney Worldwide Conservation Fund, the National Geographic Society, the Chase, Cunningham, MKCG, Offield, Peach, Thorne, Tortuga, and Kellogg Foundations, the Wadsworth Endowed Chair in Conservation Science, and Friends of the Penguins. Trimble provided field computers and Canon provided 2 pairs of binoculars. None of our funders had any influence on the content of the submitted or published manuscript. None of our funders required approval of the final manuscript to be published.

Ethics statement: This research was carried out under the University of Washington IACUC protocol #2213-02, and permits from the Office of Tourism and the Office of Flora and Fauna, Province of Chubut, Argentina. A joint agreement between the WCS and the Chubut Office of Tourism fostered the research.

Author contributions: P.D.B. designed the long-term study methods. P.D.B. and G.A.R. conducted field research and supervised field assistants. G.A.R. and P.D.B. formulated the questions. G.A.R. analyzed the data. G.A.R. wrote the paper and P.D.B. and G.A.R. edited it.

LITERATURE CITED

- Alaska Maritime National Wildlife Refuge (2016). Standardized protocols for annual seabird monitoring camps at Aiktak, Buldir, Chowiet, St. George, St. Lazaria and St. Paul islands and Cape Lisburne in the Alaska Maritime National Wildlife Refuge in 2016. U.S. Fish and Wildlife Service Report AMNWR 2016/04, U.S. Fish and Wildlife Service, Homer, AK, USA.
- Arnold, J. M., S. Brault, and J. P. Croxall (2006). Albatross populations in peril: A population trajectory for Black-browed Albatrosses at South Georgia. *Ecological Applications* 16:419–432.
- Barbraud, C., and H. Weimerskirch (2006). Antarctic birds breed later in response to climate change. *Proceedings of the National Academy of Sciences USA* 103:6248–6251.
- Beissinger, S. R., and M. I. Westphal (1998). On the use of demographic models of population viability in endangered species management. *The Journal of Wildlife Management* 62:821–841.
- Birkhead, T. R., and D. N. Nettleship (1980). Census methods for murres, *Uria* species: A unified approach. Occasional Paper 43, Canadian Wildlife Service, Ottawa, ON, Canada.
- Blight, L. K., M. C. Drever, and P. Arcese (2015). A century of change in Glaucous-winged Gull (*Larus glaucescens*) populations in a dynamic coastal environment. *The Condor: Ornithological Applications* 117:108–120.
- Boersma, P. D., and G. A. Rebstock (2009). Foraging distance affects reproductive success in Magellanic Penguins. *Marine Ecology Progress Series* 375:263–275.
- Boersma, P. D., and G. A. Rebstock (2010). Effects of double bands on Magellanic Penguins. *Journal of Field Ornithology* 81:195–205.
- Boersma, P. D., and G. A. Rebstock (2014). Climate change increases reproductive failure in Magellanic Penguins. *PLOS One* 9:e85602. doi:10.1371/journal.pone.0085602
- Boersma, P. D., and D. L. Stokes (1995). Mortality patterns, hatching asynchrony, and size asymmetry in Magellanic Penguin *Spheniscus magellanicus* chicks. In *The Penguins: Ecology and Management* (P. Dann, I. Norman, and P. Reilly, Editors). Surrey Beatty & Sons, Chipping Norton, NSW, Australia. pp. 3–25.
- Boersma, P. D., E. Frere, O. Kane, L. M. Pozzi, K. Pütz, A. Raya Rey, G. A. Rebstock, A. Simeone, J. Smith, A. Van Buren, P. Yorio, and P. García Borboroglu (2013). Magellanic Penguin (*Spheniscus magellanicus*). In *Penguins: Natural History and Conservation* (P. García Borboroglu and P. D. Boersma, Editors). University of Washington Press, Seattle, WA, USA. pp. 233–263.
- Boersma, P. D., D. L. Stokes, and P. M. Yorio (1990). Reproductive variability and historical change of Magellanic Penguins (*Spheniscus magellanicus*) at Punta Tombo, Argentina. In *Penguin Biology* (L. Davis and J. Darby, Editors). Academic Press, San Diego, CA, USA. pp. 15–43.
- Boersma, P. D., N. T. Wheelwright, M. K. Nerini, and E. S. Wheelwright (1980). The breeding biology of the Fork-tailed Storm-Petrel (*Oceanodroma furcata*). *The Auk* 97:268–282.
- Cairns, D. K. (1987). Seabirds as indicators of marine food supplies. *Biological Oceanography* 5:261–271.
- Carey, M. J. (2009). The effects of investigator disturbance on procellariiform seabirds: A review. *New Zealand Journal of Zoology* 36:367–377.
- CCAMLR (Commission for the Conservation of Antarctic Marine Living Resources) (2014). CCAMLR Ecosystem Monitoring Program: Standard Methods. Commission for the Conservation of Antarctic Marine Living Resources, North Hobart, TAS, Australia.
- Colombelli-Négrel, D. (2015). Low survival rather than breeding success explains Little Penguin population decline on Granite Island. *Marine and Freshwater Research* 66:1057–1065.
- Crawford, R. J. M., B. M. Dyer, J. Cooper, and L. G. Underhill (2006). Breeding numbers and success of *Eudyptes* penguins at Marion Island, and the influence of mass and time of arrival of adults. *CCAMLR Science* 13:175–190.
- Croxall, J. P., and P. A. Prince (1979). Antarctic seabird and seal monitoring studies. *Polar Record* 19:573–595.
- Delord, K., D. Besson, C. Barbraud, and H. Weimerskirch (2008). Population trends in a community of large procellariiforms of Indian Ocean: Potential effects of environmental and fisheries interactions. *Biological Conservation* 141:1840–1856.
- Demongin, L., M. Dmitrenok, and V. Bretagnolle (2007). Determining Great Bittern *Botaurus stellaris* laying date from egg and chick biometrics. *Bird Study* 54:54–60.
- Einoder, L. D. (2009). A review of the use of seabirds as indicators in fisheries and ecosystem management. *Fisheries Research* 95:6–13.
- Emmerson, L. M., J. Clarke, K. R. Kerry, and C. Southwell (2003). Temporal variability and the interrelationships between CEMP parameters collected on Adélie Penguins at Béchervaise Island. *CCAMLR Science* 10:75–90.

- Finkelstein, M. E., D. F. Doak, M. Nakagawa, P. R. Sievert, and J. Klavitter (2010). Assessment of demographic risk factors and management priorities: Impacts on juveniles substantially affect population viability of a long-lived seabird. *Animal Conservation* 13:148–156.
- Furness, R. W., and B. L. Furness (1981). A technique for estimating the hatching dates of eggs of unknown laying date. *Ibis* 123:98–102.
- García-Borboroglu, P., P. D. Boersma, V. Ruoppolo, L. Reyes, G. A. Rebstock, K. Griot, S. R. Heredia, A. C. Adornes, and R. P. da Silva (2006). Chronic oil pollution harms Magellanic Penguins in the Southwest Atlantic. *Marine Pollution Bulletin* 52:193–198.
- Greenwood, J. J. D., and R. A. Robinson (2006). Principles of sampling. In *Ecological Census Techniques: A Handbook* (W. J. Sutherland, Editor). Cambridge University Press, Cambridge, UK. pp. 11–86.
- Hamer, K. C., E. A. Schreiber, and J. Burger (2002). Breeding biology, life histories, and life history–environment interactions in seabirds. In *Biology of Marine Birds* (E. A. Schreiber and J. Burger, Editors). CRC Press, New York, NY, USA. pp. 217–261.
- Hodgson, J. C., S. M. Baylis, R. Mott, A. Herrod, and R. H. Clarke (2016). Precision wildlife monitoring using unmanned aerial vehicles. *Scientific Reports* 6:22574. doi:10.1038/srep22574
- Jenouvrier, S., C. Barbraud, B. Cazelles, and H. Weimerskirch (2005b). Modelling population dynamics of seabirds: Importance of the effects of climate fluctuations on breeding proportions. *Oikos* 108:511–522.
- Jenouvrier, S., C. Barbraud, and H. Weimerskirch (2005a). Long-term contrasted responses to climate of two Antarctic seabird species. *Ecology* 86:2889–2903.
- Jenouvrier, S., C. Barbraud, H. Weimerskirch, and H. Caswell (2009). Limitation of population recovery: A stochastic approach to the case of the Emperor Penguin. *Oikos* 118: 1292–1298.
- Lewis, S., D. H. Nussey, A. G. Wood, J. P. Croxall, and R. A. Phillips (2012). Intrinsic determinants of a population trend in timing of breeding in the Wandering Albatross. *Oikos* 121:2061–2071.
- Lynch, H. J., W. F. Fagan, R. Naveen, S. G. Trivelpiece, and W. Z. Trivelpiece (2009). Timing of clutch initiation in *Pygoscelis* penguins on the Antarctic Peninsula: Towards an improved understanding of off-peak census correction factors. *CCAMLR Science* 16:149–165.
- Mayfield, H. F. (1975). Suggestions for calculating nest success. *The Wilson Bulletin* 87:456–466.
- Merkel, F. R., K. L. Johansen, and A. H. Kristensen (2016). Use of time-lapse photography and digital image analysis to estimate breeding success of a cliff-nesting seabird. *Journal of Field Ornithology* 87:84–95.
- Oro, D. (2014). Seabirds and climate: Knowledge, pitfalls and opportunities. *Frontiers in Ecology and Evolution* 2:79. doi:10.3389/fevo.2014.00079
- Rebstock, G. A., P. D. Boersma, and P. García-Borboroglu (2016). Changes in habitat use and nesting density in a declining seabird colony. *Population Ecology* 58:105–119.
- Renison, D., D. Boersma, and M. B. Martella (2002). Winning and losing: Causes for variability in outcome of fights in male Magellanic Penguins (*Spheniscus magellanicus*). *Behavioral Ecology* 13:462–466.
- Reynolds, J. H., and H. M. Renner (2014). Using patch occupancy models to estimate area of crevice-nesting seabird colonies. *The Condor: Ornithological Applications* 116:316–324.
- Rolland, V., M. Nevoux, C. Barbraud, and H. Weimerskirch (2009). Respective impact of climate and fisheries on the growth of an albatross population. *Ecological Applications* 19:1336–1346.
- Sandvik, H., K. E. Erikstad, and B.-E. Sæther (2012). Climate affects seabird population dynamics both via reproduction and adult survival. *Marine Ecology Progress Series* 454:273–284.
- Schiavini, A., P. Yorio, P. Gandini, A. Raya Rey, and P. D. Boersma (2005). Los pingüinos de las costas Argentinas: estado poblacional y conservación. *Hornero* 20:5–23.
- Shaffer, T. L. (2004). A unified approach to analyzing nest success. *The Auk* 121:526–540.
- Sherley, R. B., H. Winker, R. Altwegg, C. D. van der Lingen, S. C. Votier, and R. J. M. Crawford (2015). Bottom-up effects of a no-take zone on endangered penguin demographics. *Biology Letters* 11:20150237. doi:10.1098/rsbl.2015.0237
- Southwell, C., and L. Emmerson (2015). Remotely-operating camera network expands Antarctic seabird observations of key breeding parameters for ecosystem monitoring and management. *Journal for Nature Conservation* 23:1–8.
- Southwell, C., J. McKinlay, L. Emmerson, R. Trebilco, and K. Newbery (2010). Improving estimates of Adélie Penguin breeding population size: Developing factors to adjust one-off population counts for availability bias. *CCAMLR Science* 17:229–241.
- Stokes, D. L., and P. D. Boersma (1991). Effects of substrate on the distribution of Magellanic Penguin (*Spheniscus magellanicus*) burrows. *The Auk* 108:923–933.
- Stokes, D. L., and P. D. Boersma (1998). Nest-site characteristics and reproductive success in Magellanic Penguins (*Spheniscus magellanicus*). *The Auk* 115:34–49.
- Stokes, D. L., P. D. Boersma, J. L. de Casenave, and P. García-Borboroglu (2014). Conservation of migratory Magellanic Penguins requires marine zoning. *Biological Conservation* 170:151–161.
- Walsh, P. M., D. J. Halley, M. P. Harris, A. del Nevo, I. M. W. Sim, and M. L. Tasker (1995). *Seabird Monitoring Handbook for Britain and Ireland*. JNCC/RSPB/ITE/Seabird Group, Peterborough, UK.
- Waugh, S. M., C. Barbraud, L. Adams, A. N. D. Freeman, K.-J. Wilson, G. Wood, T. J. Landers, and G. B. Baker (2015). Modeling the demography and population dynamics of a subtropical seabird, and the influence of environmental factors. *The Condor: Ornithological Applications* 117:147–164.
- Weimerskirch, H. (2002). Seabird demography and its relationship with the marine environment. In *Biology of Marine Birds* (E. A. Schreiber and J. Burger, Editors). CRC Press, New York, NY, USA. pp. 115–135.
- Wilson, P. R., D. G. Ainley, N. Nur, S. S. Jacobs, K. J. Barton, G. Ballard, and J. C. Comiso (2001). Adélie Penguin population change in the Pacific sector of Antarctica: Relation to sea-ice extent and the Antarctic Circumpolar Current. *Marine Ecology Progress Series* 213:301–309.

APPENDIX TABLE 4. Sensitivity of population growth (λ) to reproductive success is typically low in seabirds. Changes of 11% to 38% in reproductive success (or a related measure), with all other variables held constant, were required to change λ by 1% in population models for seabirds. Percent changes were reported explicitly for Adélie and African penguins and Laysan Albatrosses. We calculated the percentages for the other species based on values for reproductive success, λ , and sensitivity or elasticity reported in the papers ($\Delta RS = \Delta \lambda / S$, where ΔRS = change in reproductive success, $\Delta \lambda = 0.01$, and S = sensitivity; sensitivity = $E/(RS/\lambda)$, where E = elasticity, RS = reproductive success, and λ = population growth rate).

Species	% change in reproductive success required to change λ by 0.01	Reference
Emperor Penguin (<i>Aptenodytes forsteri</i>)	16% (Reproductive success \times 1 st -year survival)	Jenouvrier et al. (2005a)
Adélie Penguin (<i>Pygoscelis adeliae</i>)	11%	Wilson et al. (2001)
Little Penguin (<i>Eudyptula minor</i>)	19%	Colombelli-Négrel (2015)
Jackass Penguin (<i>Spheniscus demersus</i>)	20% (nestling survival)	Sherley et al. (2015)
Black-browed Albatross (<i>Thalassarche melanophris</i>)	24% (experienced breeders only)	Rolland et al. (2009)
Black-browed Albatross	38%	Arnold et al. (2006)
Laysan Albatross (<i>Phoebastria immutabilis</i>)	18%	Finkelstein et al. (2010)
Southern Fulmar (<i>Fulmarus glacialisoides</i>)	20%	Jenouvrier et al. (2005b)
Snow Petrel (<i>Pagodroma nivea</i>)	35% (Reproductive success times 1 st -year survival)	Jenouvrier et al. (2005a)