Multilevel models reveal no cohort-level variation in time spent foraging to account for a collapse in kittiwake (Rissa tridactyla) breeding success

André R. Breton\textsuperscript{a,}\ast, S. Dean Kildaw\textsuperscript{a}, K. Murra\textsuperscript{a}, C. Loren Buck\textsuperscript{b}

\textsuperscript{a} School of Fisheries and Ocean Sciences, University of Alaska Fairbanks, AK 99775-7220, USA
\textsuperscript{b} Biological Sciences Department, University of Alaska Anchorage, 3211 Providence Drive, Anchorage, AK 99508, USA

Abstract

As central-place foragers, colonial seabirds should be able to compensate, up to some threshold, for changing breeding conditions by remaining flexible in the amount of time allocated to foraging versus other activities. Black-legged kittiwake (Rissa tridactyla) colonies in Chiniak Bay, Kodiak Island, Alaska experienced high productivity from 2001 to 2003 and virtually no productivity from 2004 to 2005. In the absence of disease epidemics, increased human disturbance or predation, we applied multilevel, mixed, regression models to assess the hypothesis that the collapse from high to low breeding success was due primarily to changes in prey availability. Under this hypothesis, we predicted that longer foraging trips would be associated with reduced breeding performance in cohorts of kittiwakes—groups marked with radio-transmitters at the same colony in the same year. We separately modeled two response variables: foraging trip-durations made during the incubation and early chick stages. Multilevel mixed models revealed only weak variation at the cohort-level in either response variable; hatching success and fledging success accounted for none or less than 2% of the total variation in trip-durations made during the incubation and early chick stages, respectively. The majority of the variation in our response variables was at the observation (ca. 80–90%) and bird (or individual) levels (ca. 10–12%). Our results expose the unreliability of using indirect evidence to implicate prey availability as the primary cause of widespread breeding failures in colonial seabirds. In the ecological literature, two types of inferences seem particularly vulnerable to indirect evidence: inferences following studies on species that are accepted or strongly promoted as bioindicators of changes in marine productivity such as the black-legged kittiwake, and inferences that implicate a fisheries-induced reduction in prey availability to failed breeding. In light of these pitfalls, we recommend that long-term monitoring studies on kittiwakes and other seabirds carefully consider in their study designs and implementations multiple working hypotheses that might explain major changes in breeding success.
1. Introduction

During the breeding period, colonially nesting seabirds allocate their time between the colony, usually incubating and brooding, and the sea, mainly foraging. Flexibility in the amount of time allocated to these activities allows seabirds, as central-place foragers, to compensate for the spatial and temporal constraints associated with raising young. The idea that seabirds could compensate through flexible time budgets to dynamic food resources was discussed in a theoretical context by Cairns (1987). Since that time, several studies have shown that seabirds are able to mediate poor foraging conditions by increasing time spent foraging (Monaghan et al., 1989; Burger and Piatt, 1990; Hamer et al., 1991, 1993; Wanless and Harris, 1992; Utley et al., 1994; Suryan et al., 2002; Litzow and Piatt, 2003; Wanless et al., 2005).

Black-legged kittiwake (Rissa tridactyla; “kittiwake” hereafter) colonies in Chiniak Bay, Kodiak Island, Alaska experienced high productivity for the region (Hatch and Hatch, 1990; Roberts and Hatch, 1993; Gill and Hatch, 2002; Frederiksen et al., 2005) from 2001 to 2003 and virtually no productivity from 2004 to 2005 (Kildaw et al., 2005; Buck et al., 2007). Concurrently, intensive seabird monitoring efforts (Gulf Apex Predator–Prey Project, unpublished) did not detect evidence of disease epidemics (in any year) or an important change in human disturbance between the high and low productivity periods. And, although predation by bald eagle (Haliaeetus leucocephalus), peregrine falcon (Falco peregrinus), common raven (Corvus corax), glaucous-winged gull (Larus glaucescens) and northern gannet (Morus bassanus) on kittiwake adults, chicks, and/or eggs was observed or suspected in all years, non-systematic sampling did not detect an increase in predation after 2003. In the absence of disease epidemics, increased disturbance or predation, we applied multilevel, mixed, regression models to assess the hypothesis that the shift from high to low breeding success experienced by Chiniak Bay kittiwakes was due to a major reduction in prey productivity periods. And, although predation by bald eagle (Haliaeetus leucocephalus), peregrine falcon (Falco peregrinus), common raven (Corvus corax), glaucous-winged gull (Larus glaucescens) and northern gannet (Morus bassanus) on kittiwake adults, chicks, and/or eggs was observed or suspected in all years, non-systematic sampling did not detect an increase in predation after 2003. In the absence of disease epidemics, increased disturbance or predation, we applied multilevel, mixed, regression models to assess the hypothesis that the shift from high to low breeding success experienced by Chiniak Bay kittiwakes was due to a major reduction in prey availability. Under this hypothesis, we predicted that longer foraging trips would be associated with reduced breeding performance.

Multilevel, mixed, regression models ‘mix’ fixed and random effects in a single linear regression model (Singer, 1998; Singer and Willet, 2003). ’Multilevel’ corresponds to nesting in the response variable into one or more levels such as repeated observations nested within birds that are further nested within colonies (three levels: observation, bird, colony). Fixed or population-specific effects estimate the average population response while random or subject-specific effects estimate the subject deviations from the average population response (Singer, 1998; Fitzmaurice et al., 2004). ’Subjects’ might be observations, individuals, or groups depending on the level of the multilevel response variable. An important advantage of multilevel models is the ability to estimate random effects and their variance at each level in the response ($\sigma_{\text{random}}^2$). Subsequently, these variance components can be used to estimate the percentage of variance in $y$ accounted for by each level and, assuming the magnitude of level-i variation warrants further investigation, the percentage accounted for at level-i by a covariate.

In the simple regression model $y_{ij} = y_{00} + r_{ij}$, $y_{00}$ is the average population response (fixed intercept), $r_{ij}$ are the (assumed) random deviations from $y_{00}$ and their variance is $\sigma^2$ (notation closely follows Singer, 1998). In the context of multilevel regression, this simple model assesses variation on only one level ($i$). The variance of the $r_{ij}$ deviations $\sigma^2$ represents the residual, observation or level-1 variance component (the only variance component in the model).

Model $y_{ij} = y_{00} + \mu_{ij} + r_{ij}$ is an extension to two levels where $y_{ij}$ represents the $i$th observation from the $j$th group. As in the previous model, $y_{00}$ represents the average population response. To this model we add random group deviations from $y_{00}$ and represent these as $\mu_{ij}$; the variance of these group-level random effects is $\sigma^2_{\text{group}}$. With the inclusion of $\mu_{ij}$ in the model, the residuals take on a slightly different interpretation but their variance remains $\sigma^2$: rather than measure random deviations around the grand mean ($y_{00}$), the $r_{ij}$’s instead measure variation around the group means where $y_{ij} = y_{00} + \mu_{ij}$. Notice that if $\mu_{ij}$ is assumed to be $0$, then the $r_{ij}$’s instead measure deviation from the fixed (or population-average) level $y_{00}$ by $\mu_{ij}$.

In this model, $y_{00}$ has been ‘unfixed’ (Singer, 1998). Additional details are available in Singer (1998), Singer and Willet (2003) and Smolkowski et al. (2006).

Variance components can be used to calculate the proportion of the variance in (e.g.) $y_{ij}$ that is accounted for by the group or level-2 variance using the estimator $\hat{\sigma}_{\text{group}}^2 = \frac{1}{R \sum_{j=1}^{R} \sum_{i=1}^{N_j} (y_{ij} - \bar{y}_{ij})^2}$ where $R$ is the number of levels in $y_{ij}$. Note that if $\sigma_{\text{group}}^2$ represented only a small fraction of the variance in $y_{ij}$, then there would be no point to fitting a group-level covariate to $y_{ij}$. Alternatively, if $\sigma_{\text{group}}^2$ represented an important part of the variance, then the estimator $\hat{\sigma}_{\text{group}}^2 = \frac{1}{R \sum_{j=1}^{R} \sum_{i=1}^{N_j} (y_{ij} - \hat{Y}_{ij})^2}$ could be used to calculate the variance at level-i accounted for by a level-i covariate. Models in this equation (A, B) need to be identical except for the level-i covariate in model B (Singer, 1998; Singer and Willet, 2003).

We used mixed models with up to four levels to estimate variance components in the responses $y_{\text{incubation}}$ and $y_{\text{stage}}$ where $y_{\text{stage}}$ describes the duration (h) of the ith foraging trip (or observation; level-1) from the jth bird (or individual; level-2), kth cohort (level-3) and lth colony (level-4) measured during the incubation or the early chick stage. Cohorts are groups of breeding adults fitted with radio-transmitters at the same colony in the same year. Conditional on the assumption that foraging trip-duration is a reliable proxy of moderate to large changes in prey availability (Cairns, 1987; Monaghan et al., 1989; Burger and Piatt, 1990; Hamer et al., 1991, 1993; Wanless and Harris, 1992; Utley et al., 1994; Suryan et al., 2002; Litzow and Piatt, 2003; Wanless et al., 2005), we predicted that if food available to kittiwakes plummeted after 2003, then breeding success, as a proxy of prey abundance, would account for an important amount of the cohort-level variation in time spent foraging. In particular, we expected that foraging trip–durations made by cohorts marked...
in 2004 and 2005 (low productivity years) would be (on average) substantially longer than those marked from 2001 to 2003 (high productivity years). It is important to note that if estimates of prey abundance had been available, our analysis would have looked for a cause and effect relationship between prey availability and trip-duration. However, given that estimates of prey abundance were unavailable, we employed what is widely believed to be a proxy of prey availability, breeding success, and looked for a correlation (not cause and effect) between breeding success and foraging trip-duration.

2. Methods

2.1. Study area and species

Chiniak Bay, Kodiak Island, Alaska (57°40'N, 152°20'W) is ca. 400 km² and contains 21 kittiwake colonies encompassing ca. 10,000 breeding pairs from about April to August (breeding period) each year (Kildaw et al., 2005). Colonies are on cliffs throughout the main bay and three secondary bays on numerous islands, sea stacks, and on Kodiak Island itself (see Fig. 1 in Kildaw et al., 2005). Data for our analyses were collected at five island colonies (Gull, Kalsin, Kulichkof, Mary, and Svitlak) and two mainland sites (Gibson Cove and Sealand). The mainland sites are adjacent (see Fig. 1 in Kildaw et al., 2005) so were combined in our analyses and presented below as Gibson Cove.

Black-legged kittiwakes feed by plunge diving from a few meters above the sea surface to a depth of 0.5–1.0 m, or surface seizing while sitting on the water (Baird, 1994). The principle prey of breeding kittiwakes in the Kodiak region are Pacific sand lance (Ammodytes hexapterus) and capelin (Mallotus villosus), both surface-schooling fishes (Baird, 1990; Kildaw et al., 2005). Like other members of the family Laridae, kittiwakes tend to nest colonially, are sexually dimorphic (males often larger than females) and generally monogamous (Coulson and Thomas, 1985). The typical clutch for the species is two eggs, although single egg clutches are not uncommon (range 1–3; Coulson and Thomas, 1985; Murphy et al., 1991). Incubation and brood rearing occur over 25–29 and 34–58 days, respectively (Maunnder and Threlfall, 1972; Baird, 1994; Gill et al., 2002). The latter period is often divided into early and late chick stages in analyses similar to our own (e.g., Gill et al., 2002).

2.2. Radio-telemetry and foraging trips

Adult kittiwakes were captured with noose poles or foot snares at their nests, measured and fitted with radio-transmitters (n = 183). Birds were captured during both incubation and early chick stages (see below). In three instances, the same bird was captured and deployed with a transmitter in two separate years. In each case, we included data from only one of the years (the year with the fewest number of marked birds). Data from 12 individuals came from their second rather than first within-season nesting attempt. At all nests except one, only one adult was marked with a transmitter; from the pair, we excluded the female since fewer males were available (see Section 3).

Nests of radio-marked birds were checked every 5 days with few exceptions to assess nest status. We excluded foraging trips from our analyses that occurred after nest failure and from periods when the egg(s) were known or suspected to be addled. Date of failure was recorded as the last day that an egg or chick was confirmed present. Since no birds were marked with radio-transmitters in the first 2 weeks of incubation, our incubation stage represents the second half of this period. Our early chick stage represents the period of rapid growth from hatch (day 0) to the approximate end of the linear growth phase, day 20 (Maunnder and Threlfall, 1972; Baird and Gould, 1986).

Transmitters (Advanced Telemetry Systems (ATS), Isanti, Minnesota, USA) emitted a unique frequency in the 166–168 MHz range and weighed ≤12 g representing ca. 3% of the average mass of kittiwakes in this study (400.85 ± 33.15 S.D.). Transmitters were attached parallel to the underside of three central tail feathers and close to the body with a quick-setting glue (Loctite®422) and accelerometer (Loctite®712). All procedures were approved by the University of Alaska Fairbanks Institutional Animal Care and Use Committee (IACUC #02-12, 05-43) and authorized under state and federal permits.

Automated, telemetry, receiving-stations (ATS, Isanti, Minnesota, USA) were deployed at each colony and consisted of a model R4000 receiver connected to a coaxial cable antenna and a data collection computer (DCC II and Programming DCC). Systems were powered with a 30 Ah gel-cell battery maintained with a 45 W solar panel connected to a 5 A charge regulator. In 2001, the detection range of the receiver was limited to ca. 2 km or less from the colony. In 2002–2005 attenuators were used to restrict the detection range to less than 300 m from the colony. This assured that detections were from birds within or adjacent to the colony. The detection range of each system was determined by affixing a transmitter to a 3 m pole on a boat, running 1–3 linear transects away from the colony and determining at what point the signal was lost at the colony.

Radio-telemetry systems were programmed to cycle through frequencies every ≤10 min. When a frequency was detected, the system stopped cycling for 5 s and logged the date, time, frequency and pulse count before resuming cycling through frequencies. Radio-transmitters emitted 45–50 pulses min⁻¹; therefore, a normal pulse count was ca. 3–5 pulses. Pulse counts of ca. 6–11 were possible if birds were very close to the antenna and strong signals were double counted. We visited systems every 5–10 days to upload data from the DCCs to a laptop computer.

We processed radio-telemetry data using a custom program that output start-time, stop-time and duration of each bout of presence (within range of the colony telemetry station) and absence (out of range of the colony telemetry station). Although bouts of absence beyond the detection area of the colony-based telemetry systems included activities other than time spent foraging (e.g., commuting), we assumed that time spent away from the colony was a reliable proxy of time spent foraging (Suryan et al., 2000).

We required birds to be detected three times within a 45 min period before we considered them present and they remained in this state until ≥45 min elapsed without a detection. At this time, the presence bout was terminated and a bout...
of absence (foraging trip) was initiated. To minimize bias in foraging trip-durations produced by false detections (noise), we scanned the dataset visually. Signal noise generally occurred in conspicuous blocks and was indicated by detections of dummy frequencies (known to be inactive) or by detections of deployed frequencies with pulse counts outside of the 3–11 pulse range (radio-frequency static(interference typically produces high pulse counts in the 15–100+ range). To avoid bias from data gaps that resulted from discarded blocks of noise, system failures and other problems, any foraging trip that was preceded or followed by a system interruption of ≥45 min was excluded from our analyses.

2.3. Productivity plots

Productivity plots contained ca. 30 nests and were monitored at all colonies to calculate annual hatching and fledging success—breeding success covariates that overlap temporally with incubation and early chick stage foraging trips, respectively. The number of plots at each colony (μ = 5.4; range 1–11) was a function of colony size and accessibility; large colonies generally had more plots than small colonies. Hatching and fledging success were calculated as eggs hatched/eggs laid and chicks fledged/chicks hatched, respectively for each plot and then averaged across plots for each colony. A chick was considered to have fledged if its estimated age was 36 days when it was no longer present at the nest; however, chicks were assumed to have survived if present on the last visit of the season and were at least 25 days old.

2.4. Observed cohort means

As part of our analyses, mean foraging trip-durations (±S.E.) for each cohort during both the incubation and early chick stage were plotted with hatching (incubation stage) and fledging success (early chick stage) against year, respectively. Mean trip-durations by cohort were estimated using mean trip-durations from each bird. Plots of breeding success provide an initial indication of variation among cohorts and the correlation between trip-duration and our measures of breeding success.

2.5. Statistical analyses

We used mixed regression models with up to four levels to estimate variance components in the responses \( y_{ijkl} \) and \( y_{ijkl}^{\text{early chick}} \) (Singer, 1998; Smolkowski et al., 2006). \( y_{ijkl}^{\text{stage}} \) describes the duration (h) of the ith foraging trip (or observation; level-1) from the jth bird (level-2), kth cohort (level-3) and lth colony (level-4). Foraging trips from each stage were analyzed separately. Eight unconditional models (no covariates; unconditional means models in Singer, 1998) were initially fit to each dataset and ranked using AICc (see below). These unconditional means models in Singer, 1998 were incorporated into two covariate models: a model with breeding success (hatching or fledging success) and another with breeding success, clutch size, and sex. The latter model controls for potential confounding by clutch size and sex (level-2 covariates). Variance components from an unconditional model (all AICc supported random effects) and each conditional model were used to estimate the proportion of cohort variation accounted for by our breeding success covariates. If present, covariate models were ranked using AICc, with the unconditional models.

Models were fitted to our response variables using PROC MIXED in SAS 9.1.3 (Singer, 1998). In order to compare models without the same fixed effects, we specified the full information maximum likelihood (FIML) estimation method in all models (Singer, 1998; Smolkowski et al., 2006). Although FIML variance components estimates can be biased, this is minimized when models contain few fixed effects relative to groups (Singer, 1998; Smolkowski et al., 2006). Our models contained 16 cohorts and <4 fixed effects. Residual plots showed that our response variables were strongly right skewed. Residual and normal probability plots of the log-transformed data did not reveal important deviations from normality—here we report analyses of these log-transformed data.

Fixed effects in our models were an intercept \( (\mu_{0000}) \), clutch size \( (\mu_{0010}^{\text{Clutch size}}) \), sex \( (\mu_{0020}^{\text{Sex}}) \), and a measure of breeding success \( (\mu_{0020}^{\text{BS}}) \) (notation closely follows Singer, 1998). Subscripting identifies the level of the effect. For example, \( \mu_{0020}^{\text{Sex}} \) refers to the second level-2 fixed-effect, sex. Sex was determined either by extraction of DNA from whole blood using DNEasy Extraction Kits (QIAGEN Inc., Valencia, CA), PCR amplification and visualization of PCR products following Griffiths et al. (1998); or published discriminant functions (Jodice et al., 2000). Blood samples and measurements were collected when radio-transmitters were deployed.

The level-1 random effect \( r_{ijkl} \) was combined with up to three additional random effects: \( \mu_{ijkl}^{\text{Sex}} \) and \( \mu_{ijkl}^{\text{Clutch size}} \), and \( \mu_{ijkl}^{\text{BS}} \) are level 2–4 random intercepts, respectively. We assumed that each set of random effects, e.g., the set of cohort-level random effects (one for each year and colony combination), were normally distributed with mean zero and variance \( \sigma^2 \). \( r_{ijkl}^{\text{level}} \sim N(0, \sigma^2) \), \( \mu_{ijkl}^{\text{Sex}} \sim N(0, \sigma^2_{\text{Sex}}) \), \( \mu_{ijkl}^{\text{Clutch size}} \sim N(0, \sigma^2_{\text{Clutch size}}) \), and \( \mu_{ijkl}^{\text{BS}} \sim N(0, \sigma^2_{\text{BS}}) \). Although our interest and focus is biological variation, these variance components include an unknown quantity of measurement error.

Models were ranked from best to worst in each set (incubation, early chick) using Akaike’s information criterion (AIC) adjusted for small sample size (AICc). AIC attempts to identify the amount of Kullback–Leibler (K-L) information lost by a model relative to truth (Kullback and Leibler, 1951; Akaike, 1973). Heuristically, AIC can be interpreted as the distance between each model and truth (Burnham and Anderson, 2002). We calculated AICc differences between the highest ranking model (lowest AICc score) and all other models in the set, and Akaike weights (Burnham and Anderson, 2002). To assist in identifying important variance components and covariates, we conservatively defined top models as those within 6 AICc.

---

1 Following our analyses, FIML and REML (restricted maximum likelihood) variance estimates were compared and found to be essentially identical.
units from the best model (many authors use a difference of 2 but this is flawed; see pages 70–79 in Burnham and Anderson, 2002). Akaike weights ($w_i$) can be interpreted as the probability that model $q$ is the best K-L model in the set (Burnham and Anderson, 2004). And since the weights sum to one across the full model set, these can be used to judge the relative support for one model or a set of models.

3. Results

3.1. Sex

Of the 108 birds in our analyses, the sexes of 37 and 16 were discriminated using head-to-bill (HB) with flattened wing (FW) and HB alone, respectively (models 6 and 7 in Table 3, Jodice et al., 2000). Accuracy predictions for discriminant functions used were 92% (HB + FW) and 88% (HB); prediction-DEE birds in Table 4, Jodice et al., 2000). The remaining 55 birds were sexed using DNA. From these methods, 69 and 39 were identified as females and males, respectively (Table 1).

3.2. Plots of means and breeding success

Two broad patterns are apparent in our plots of observed cohort means and breeding success (Figs. 1 and 2): there is consistency in the mean duration of foraging trips among cohorts and colonies; no relationship between foraging trip-duration and our metrics of breeding success is apparent. Based on these observations, we predicted that cohort- and colony-level variation in foraging trip-durations would represent a small fraction of the variance in $\gamma_{ijkl}$ inc and $\gamma_{ijkl}$ early chick.

3.3. Multilevel, mixed, regression models

One extreme outlier was removed leaving 923 incubation and 1755 early chick stage foraging trips for our analyses. In the incubation model set (Table 2), cohort ($\theta_{2(cohort)}$) and colony ($\theta_{2(colony)}$) variance components were too small to be estimated in most models. Colony was estimable in one incubation model but this model acquired no support (\Delta AICc = 33.7; Table 2). In models where $\theta_{2(cohort)}$ was estimable (Table 2), the percentage of the variance in incubation foraging trip-durations ($\gamma_{ijkl}$ incubation) that was accounted for by the cohort-level variance was less than 1/1000th of one percent (Table 3). About 10.5% of the variance in $\gamma_{ijkl}$ incubation was accounted for by bird-level variation ($\theta_{2(bird)}$; Table 3). Random colony and cohort effects were too small to be estimated in the full random effects model. Given that there was apparently no cohort variation in the data to be accounted for by a cohort-level covariate, we did not fit any covariate models to incubation foraging trips.

In the early chick model set, all fixed and random effects were estimable in all models. Three unconditional models (no covariates) acquired 87% of the AICc weight. We therefore excluded model $\mu_{ijkl} + \theta_{ijkl}$ which added only an additional 3% of the AICc weight (Table 4). Given AICc support for bird, cohort, and colony variance components (Table 4), covariates were fitted to model $\mu_{ijkl} + \theta_{ijkl} + \theta_{ijkl}$ (includes all random effects). Inclusion of our covariates did not improve model fit: fit of covariate models was poorer than the reference model, $\mu_{ijkl} + \theta_{ijkl} + \theta_{ijkl}$; the full covariate model was not supported (\Delta AICc = 6.8; Table 4).

In models including all random effects, bird, cohort ($\theta_{2(cohort)}$) and colony ($\theta_{2(colony)}$) variance components accounted for about 10%, 4%, and 4% of the variance in early chick trip-durations ($\gamma_{ijkl}$ early chick), respectively (Table 5). As in the incubation model set, there was strong evidence that variation among birds was an important component of the variance: all top models included bird-level variance (Table 4) and the bird-level variance component in each model was moderately precise (Table 5). Although cohort and colony percentages were small, \Delta AICc suggests that a component of variance for colony and cohort levels is highly probable given the data. However, $\theta_{2(cohort)}$ and $\theta_{2(colony)}$ standard errors were typically larger than the estimates themselves demonstrating a high degree of uncertainty in the size of these variance components.

The conditional and unconditional models, $\gamma_{ijkl}(FS_{ij}) + \mu_{ijkl} + \theta_{ijkl} + \theta_{ijkl}$ and $\mu_{ijkl} + \theta_{ijkl} + \theta_{ijkl}$, respectively, were used to estimate the percentage of cohort

---

**Table 1 – Sample sizes of individuals fitted with radio-transmitters at six colonies in Chiniak Bay, Kodiak Island, Alaska, used in our analyses and grouped by colony, year, and sex (females/males)**

<table>
<thead>
<tr>
<th>Colony</th>
<th>2001</th>
<th>2002</th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
<th>Colony totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gibson Cove</td>
<td>2/3</td>
<td>0</td>
<td>0</td>
<td>1/1</td>
<td>3/2</td>
<td>6/6</td>
</tr>
<tr>
<td>Gull Island</td>
<td>0</td>
<td>6/2</td>
<td>8/4</td>
<td>2/2</td>
<td>1/3</td>
<td>17/11</td>
</tr>
<tr>
<td>Kalsin Island</td>
<td>5/0</td>
<td>5/0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10/0</td>
</tr>
<tr>
<td>Kulichof Island</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5/2</td>
<td>5/2</td>
<td>5/2</td>
</tr>
<tr>
<td>Mary Island</td>
<td>3/2</td>
<td>11/3</td>
<td>3/5</td>
<td>0</td>
<td>1/0</td>
<td>18/10</td>
</tr>
<tr>
<td>Svitlak Island</td>
<td>0</td>
<td>2/1</td>
<td>2/4</td>
<td>5/4</td>
<td>4/1</td>
<td>13/10</td>
</tr>
<tr>
<td>Year totals</td>
<td>9/5</td>
<td>24/6</td>
<td>13/13</td>
<td>8/7</td>
<td>14/8</td>
<td>69/39</td>
</tr>
</tbody>
</table>

a Adjacent Gibson Cove and Seal Island colonies were combined in our analyses; they are located on a cliff face attached to Kodiak Island; all others are on islands within Chiniak Bay (see Fig. 1 in Kildaw et al., 2005).

b Total for all years and colonies.

---

**Table 2 – Model set used to assess the importance of observation-, bird-, cohort-, and colony-level variance components in foraging trips made during the incubation stage**

<table>
<thead>
<tr>
<th>Model $^a$</th>
<th>Random effects $^b$</th>
<th>$\Delta$AICc $^b$</th>
<th>AICc $^b$</th>
<th>$w_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu_{ijkl} + \gamma_{ijkl}$</td>
<td>Bird, observation</td>
<td>0.0</td>
<td>0.731</td>
<td></td>
</tr>
<tr>
<td>$\mu_{ijkl} + \mu_{ijkl} + \gamma_{ijkl}$</td>
<td>Bird, cohort, observation</td>
<td>2.0</td>
<td>0.269</td>
<td></td>
</tr>
<tr>
<td>$\mu_{ijkl} + \gamma_{ijkl}$</td>
<td>Cohort, observation</td>
<td>30.1</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>$\gamma_{ijkl}$</td>
<td>Observation</td>
<td>32.5</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>$\mu_{ijkl} + \gamma_{ijkl}$</td>
<td>Colony, observation</td>
<td>33.7</td>
<td>0.000</td>
<td></td>
</tr>
</tbody>
</table>

All parameters were estimable in these models. Models with non-estimable parameters are not shown; see Section 3 for details.

$a$ A fixed intercept $\mu_{ijkl}$ is present in all models but not shown here; $\mu_{ijkl}$, $\mu_{ijkl}$ and $\gamma_{ijkl}$ are random effects and their variance components are $\theta_{2(bird)}$, $\theta_{2(cohort)}$, $\theta_{2(colony)}$ and $\theta_{ijkl}$, respectively; observation represents the residual ($\gamma$) variation.

$b$ AICc for the best model (AICc = 0.00) was 2530.8.
Fig. 1 – (a–f) Mean foraging trip-duration in hours (left axis, bars) during incubation and hatching success (right axis, solid circles) against year for all cohorts (±S.E.). When bars are absent, foraging data for these cohorts were not available. Numbers represent the number of birds (bottom-left corner, bars) and productivity plots used to estimate mean trip-durations and fledging success, respectively.

4. Discussion

We monitored six black-legged kittiwake colonies distributed across a ca. 400 km² bay over a 5-year period. During this time, breeding success was initially high relative to the extremely poor 2004 and 2005 seasons when ca. 10,000 kittiwake pairs produced less than 1/10th of a chick pair⁻¹. Given the extent of the breeding failure, and an absence of disease epidemics, increased human disturbance or predation, we suspected a substantial decline in prey availability to be the root cause. However, in our plots of observed cohort means, changes in the mean duration of foraging trips did not show a clear relationship to variation in hatching or fledging success during incubation and early chick stages, respectively; mean trip-durations were fairly consistent across cohorts. As predicted by these plots, unconditional, multilevel, mixed models revealed very little variation in foraging trip-durations at the cohort-level. Hatching success and fledging success accounted for none or <2% of this cohort-level variation, respectively. The majority of the variation in incubation and early chick forag-
ing trip-durations was at the observation (ca. 80–90%) and bird levels (ca. 10–12%). Although the temporal overlap between trip-duration, hatching and fledging success are far from perfect, these results demonstrate a clear disconnect between these variables.

Under the assumptions that (1) moderate to large changes in prey availability will trigger similar changes in time spent foraging (see references in Section 1) and (2) our response variables accurately measured time spent foraging (Wanless, 1992; Suryan et al., 2000), our results refute the hypothesis that a major change in food availability alone (during the stages that we monitored: late incubation; early chick) led to poor breeding success in 2004 and 2005. Instead, our results implicate some other factor(s) working alone or in conjunction with prey availability or prey quality. We propose four alternative hypotheses to explain the collapse in kittiwake breeding success and the lack of cohort-variation in our response variables: Chiniak Bay kittiwakes might be foraging at maximum capacity in good years which would preclude them from buffering reduced (1) prey availability or (2) prey quality by increasing time spent foraging (Cairns, 1987; Burger and Piatt, 1990; Litzow and Piatt, 2003; Wanless et al., 2005); (3) low breeding success in 2004 and 2005 may be due to an interaction between predation and prey availability or quality (two scenarios are considered; Hatch and Hatch, 1990; Regehr and Montvecchi, 1997).

Fig. 2 – (a–f) Mean foraging trip-duration in hours (left axis, bars) during the early chick stage and fledging success (right axis, solid circles) against year for all cohorts (±S.E.). When bars are absent, foraging data for these cohorts were not available. Numbers represent the number of birds (bottom-left corner, bars) and productivity plots used to estimate mean trip-durations and fledging success, respectively.
Evidence in support of this buffering hypothesis increased foraging time (Burger and Piatt, 1990; Litzow and Hatch, 2002; Frederiksen et al., 2005). From a controlled experiment on Middleton Island in the Gulf of Alaska, Gill and Hatch (2002) demonstrated that kittiwake productivity could be increased through supplemental feeding. At nests that were fed throughout the breeding season compared to controls (no supplemental feeding), overall productivity (chicks fledged/sites with pairs) increased by over 0.5 chick pair\(^{-1}\) in both years of the study.

Underlying Cairns’ (1987) arguments linking time spent foraging to prey availability is the idea that central-place foragers have the option to “buffer” poor feeding conditions during the breeding period through reduced colony attendance and aging for prey availability is the idea that central-place foragers have historically been much less productive than conspecifics breeding in the North Atlantic (Hatch and Hatch, 1990; Gill and Hatch, 2002; Frederiksen et al., 2005). From a controlled experiment on Middleton Island in the Gulf of Alaska, Gill and Hatch (2002) demonstrated that kittiwake productivity could be increased through supplemental feeding. At nests that were fed throughout the breeding season compared to controls (no supplemental feeding), overall productivity (chicks fledged/sites with pairs) increased by over 0.5 chick pair\(^{-1}\) in both years of the study.

If food supply was low enough to reduce chick production by over 0.5 chick pair\(^{-1}\) in the adjacent Gulf of Alaska, then kittiwakes in our study may have been feeding at maximum capacity in the good reproductive years, 2001–2003 and hence unable (physiological constraints) or unwilling (maximize lifetime reproduction; Stearns, 1992) to buffer a reduction in prey availability by increasing time spent foraging in the poor reproductive years. This would explain why our analyses failed to detect important cohort-level variation in time spent foraging during, in particular, the early chick stage—the most energetically demanding stage of the reproductive period in kittiwakes (Moe et al., 2002).

Similarly, a reduction in prey quality after 2003 could also explain failed breeding in 2004–2005. Prey quality was recently implicated as the proximate cause of a major breeding failure at the Isle of May, Scotland, an important seabird breeding site in the North Sea (Wanless et al., 2005). In a separate manuscript using extensive prey samples collected from

### Table 3 – Estimates of variance components ± S.E. from top AIC\(_C\) models (\(\Delta\text{AIC}_C \leq 6\); Table 2) fitted to incubation stage foraging trips

| Variance components | Models, random effects (\(\mu, r\))
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(\tilde{\sigma}_1^2 \pm \text{S.E.})</td>
<td>(\mu_0jkl + r_{ijkl}) (cohort, obs.)</td>
</tr>
<tr>
<td>(\tilde{\sigma}_2^2 \pm \text{S.E.})</td>
<td>0.844 (0.041)</td>
</tr>
<tr>
<td>(\tilde{\sigma}_3^2 \pm \text{S.E.})</td>
<td>0.098 (0.031)</td>
</tr>
<tr>
<td>(\sum \tilde{\sigma}_{ijkl}^2 / \text{level-1} \times 100)</td>
<td>10.43%</td>
</tr>
<tr>
<td>(\sum \tilde{\sigma}_{ijkl}^2 / \text{level-1} \times 100)</td>
<td>0.07 \times 10(^{-2})%</td>
</tr>
</tbody>
</table>

Also shown are the percentages of variation in foraging trips accounted for by bird-level (\(\tilde{\sigma}_2^2\)) and cohort-level (\(\tilde{\sigma}_3^2\)) variance components.

* obs. = residual (r) variation or variation at the observation level (level-1).

### Table 4 – Model set used to assess the importance of observation-, bird-, cohort-, and colony-level variance components and three covariates (fixed effects) in foraging trips made during the early chick stage

<table>
<thead>
<tr>
<th>Model*</th>
<th>Fixed effects (level)*</th>
<th>Random effects*</th>
<th>(\Delta\text{AIC}_C)</th>
<th>\text{AIC}_Cw_1</th>
</tr>
</thead>
<tbody>
<tr>
<td>(H_{ijkl} + \mu_{ijkl} + r_{ijkl})</td>
<td>Bird, cohort, observation</td>
<td>0.0</td>
<td>0.358</td>
<td></td>
</tr>
<tr>
<td>(H_{ijkl} + \mu_{0ijkl} + r_{ijkl})</td>
<td>Bird, colony, observation</td>
<td>0.5</td>
<td>0.279</td>
<td></td>
</tr>
<tr>
<td>(H_{ijkl} + \mu_{ijkl} + \mu_{0ijkl} + r_{ijkl})</td>
<td>Bird, cohort, colony observation</td>
<td>0.9</td>
<td>0.228</td>
<td></td>
</tr>
<tr>
<td>(\gamma_{0010}(\text{FS}<em>{ijkl}) + H</em>{ijkl} + \mu_{ijkl} + \mu_{0ijkl} + r_{ijkl})</td>
<td>Fledging success (3)</td>
<td>2.8</td>
<td>0.088</td>
<td></td>
</tr>
<tr>
<td>(H_{ijkl} + r_{ijkl})</td>
<td>Bird, observation</td>
<td>5.1</td>
<td>0.028</td>
<td></td>
</tr>
<tr>
<td>(\gamma_{0010}(\text{Clutch}<em>{ijkl}) + \gamma</em>{ijkl}(\text{Sex}<em>{ijkl}) + \gamma</em>{ijkl}(\text{FS}<em>{ijkl}) + H</em>{ijkl} + \mu_{ijkl} + \mu_{0ijkl} + r_{ijkl})</td>
<td>Bird, cohort, colony observation</td>
<td>6.8</td>
<td>0.012</td>
<td></td>
</tr>
<tr>
<td>(\mu_{ijkl} + r_{ijkl})</td>
<td>Cohort, observation</td>
<td>78.6</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>(\mu_{0ijkl} + \mu_{00jkl} + r_{ijkl})</td>
<td>Cohort, colony, observation</td>
<td>80.1</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>(H_{ijkl} + r_{ijkl})</td>
<td>Colony, observation</td>
<td>125.4</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>(r_{ijkl})</td>
<td>Observation</td>
<td>197.4</td>
<td>0.000</td>
<td></td>
</tr>
</tbody>
</table>

All parameters were estimable in these models.

* A fixed intercept \(\gamma_{0000}\) is present in all models but not shown here; \(H_{ijkl}, \mu_{ijkl}, \mu_{0ijkl}\) and \(r_{ijkl}\) are random effects and their variance components are \(\tilde{\sigma}_1^2, \tilde{\sigma}_2^2, \tilde{\sigma}_3^2\), and \(\tilde{\sigma}_4^2\), respectively; observation represents the residual (r) variation. \(\gamma_{0010}(\text{FS}_{ijkl}), \gamma_{ijkl}(\text{Clutch}_{ijkl})\) and \(\gamma_{ijkl}(\text{Sex}_{ijkl})\) are fixed effects representing fledging success, clutch size, and sex covariates, respectively. Numbers in parentheses identify the level to which the fixed effect was applied.

* AIC\(_C\) for the best model (\(\Delta\text{AIC}_C = 0.00\)) was 3949.2.
Table 5 – Estimates of variance components ± S.E. from top AICc models (ΔAICc ≤ 6; Table 4) fitted to early chick stage foraging bouts

<table>
<thead>
<tr>
<th>Variance components</th>
<th>Models, fixed (γ) and random effects (μ, r)a</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>μ₀jkl + μ₀₀00 + rjkl (bird, cohort, obs.)</td>
</tr>
<tr>
<td>δ² ± S.E.a</td>
<td>0.522 (0.018)</td>
</tr>
<tr>
<td>δ²_bird ± S.E.</td>
<td>0.062 (0.018)</td>
</tr>
<tr>
<td>δ²_cohort ± S.E.</td>
<td>0.055 (0.034)</td>
</tr>
<tr>
<td>δ²_colony ± S.E.</td>
<td></td>
</tr>
</tbody>
</table>

∑E=E EA-1 level-i × 100

δ²_bird accounted for by FSb

9.7 12.5 10.0 10.1

δ²_cohort accounted for by FSc

8.5 4.1 4.1 4.1

δ²_colony accounted for by FS

4.5 4.0 3.8

Also shown are the percentages of variation in foraging trips accounted for by bird-level (δ²_bird), cohort-level (δ²_cohort) and colony-level (δ²_colony) variance components and the percentage of cohort variation accounted for by fledging success (FS).

<table>
<thead>
<tr>
<th>Variance components</th>
<th>Models, fixed (γ) and random effects (μ, r)a</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>μ₀jkl + μ₀₀00 + rjkl (bird, cohort, obs.)</td>
</tr>
<tr>
<td>δ² ± S.E.a</td>
<td>0.522 (0.018)</td>
</tr>
<tr>
<td>δ²_bird ± S.E.</td>
<td>0.062 (0.018)</td>
</tr>
<tr>
<td>δ²_cohort ± S.E.</td>
<td>0.055 (0.034)</td>
</tr>
<tr>
<td>δ²_colony ± S.E.</td>
<td></td>
</tr>
</tbody>
</table>

∑E=E EA-1 level-i × 100

δ²_bird accounted for by FSb

9.7 12.5 10.0 10.1

δ²_cohort accounted for by FSc

8.5 4.1 4.1 4.1

δ²_colony accounted for by FS

4.5 4.0 3.8

Also shown are the percentages of variation in foraging trips accounted for by bird-level (δ²_bird), cohort-level (δ²_cohort) and colony-level (δ²_colony) variance components and the percentage of cohort variation accounted for by fledging success (FS).

a obs. = residual (r) variation or variation at the observation level (level-1).

b Higher precision was necessary to show the decrease in cohort variance among models.

c The estimator (δ²_group,A - δ²_group,B/δ²_group,A) × 100 with models μ₀jkl + μ₀₀00 + μ₀₀₀0 + rjkl (A) and γ₀₀₀₁(FSjk) + μ₀₀₀₀ + μ₀₀₀0 + μ₀₀₀0 + rjkl (B) was used to estimate cohort variation accounted for by fledging success (FS).
Chiniak Bay kittiwakes from 2001 to 2005 (C.L. Buck, unpublished), we intend to explore the possibility that prey quality declined after 2003 and subsequently led to failed breeding.

Glaucous-winged gulls and kittiwakes nest sympatrically in Chiniak Bay and rely on the same prey resources to feed their chicks (primarily sand lance and capelin; C.L. Buck unpublished). Given this overlap, there is an opportunity for increased predation by gulls on kittiwakes when the abundance or quality of sand lance and capelin decline (Regehr and Montevecchi, 1997). Although breeding success by glaucous-winged gulls in 2004 and 2005 was essentially zero, we did not observe a single instance of depredation by glaucous-winged gulls on kittiwake eggs or chicks in any year. In our view, this effectively excludes the possibility that depredation by glaucous-winged gulls played an important role in the collapse in kittiwake breeding success. However, depredation by bald eagles, peregrine falcons and common ravens may have increased in the poor breeding years following a weakening in resistance to forced predation.

At least two forms of nest predation have been proposed, opportunistic and forced. The former occurs when nests are left unguarded and predators opportunistically depredate eggs and chicks (Hatch and Hatch, 1990). The latter requires the predator to force the adult(s) from the nest prior to depredating nest contents (Regehr and Montevecchi, 1997). One obvious difference between forced and opportunistic nest predation is the presence of a nest-attending-adult in the forced case. A subtle difference is that forced predation can be met with a range of resistance including: no resistance, the predator drives off the adult by just approaching the nest; or strong resistance, the adult defends the nest until some physical threshold is reached (e.g., exhaustion). It follows then that strong colony attendance by adults is not mutu-
ally exclusive to substantial losses of eggs and chicks to nest predators.

Kittiwakes in our study did not show an obvious increase in nest abandonment from the high to low productivity years and colony attendance was strong in all years. Therefore, a high loss of eggs and chicks to opportunistic nest predation is an unlikely contributor to the collapse in productivity. However, a weakened response to forced nest predation by peregrine falcons and common ravens could have led to high egg and chick loss; these species were frequently implicated in depredation events observed or suspected in Chiniak Bay. Consistent with our results, we would not expect forced predation to cause an increase in foraging trip-durations in 2004–2005. Underlying the weakened response could have been a reduction in prey quality after 2003 which jeopardized the body condition of breeding kittiwakes thereby increasing corticosterone levels and reducing resistance to forced predation (Wingfield and Sapolsky, 2003). Interestingly, Buck et al. (2007) found that baseline corticosterone levels were significantly ($p < 0.05$) higher in 2004–2005 relative to 2001–2003 in breeding Chiniak Bay kittiwakes. However, they also found that variation in adult body condition among years was not significantly different ($p < 0.05$).

High dependence on marine productivity in northern temperate and polar regions (Nettleship, 1991) makes seabirds particularly susceptible to disruptions in warm and cold seawater movements; these shifts are directly linked to climate warming and cooling (Behrenfeld et al., 2006). Given ongoing and predicted atmospheric warming, regional and larger scale reductions in primary production are probably already underway in some areas (e.g., see Wanless et al., 2005). We speculate that as climate forcing plays out, reduced prey quality will precede reductions in prey availability. Concerning Pacific black-legged kittiwakes, indirect evidence pointing to prey availability as the source of reduced breeding (e.g., Hatch and Hatch, 1990; Baird, 1990; Murphy et al., 1991) is now supported by the controlled supplemental feeding study reported by Gill and Hatch (2002). However, none of these studies preclude prey quality as a plausible competing hypothesis.

Our results expose the unreliability of using indirect evidence to implicate prey availability (during the stages that we monitored: late incubation; early chick) as the primary cause of widespread breeding failures in colonial seabirds. Our study is a case in point whereby we used indirect measures of prey availability (e.g., no change in predation after 2003 and near complete breeding failure in 2003 and 2004) to invoke the hypothesis that prey availability crashed after 2003 and led to a collapse in breeding success. In the light of our results, we can now accept that our indirect measures were not reliable indicators of changes in prey availability per se—i.e., they likely conceal two or more major factors which may or may not include prey availability.

In the ecological literature, two types of inferences seem particularly vulnerable to indirect evidence: inferences following studies on species that are accepted or strongly promoted as bioindicators of changes in marine productivity (see also Regehr and Montevecchi, 1997) such as the black-legged kittiwake (Harris and Wanless, 1990; Gill et al., 2002); and inferences that implicate a fisheries-induced reduction in prey availability to failed breeding. In light of these pitfalls, we recommend that long-term monitoring studies on kittiwakes and other seabirds carefully consider in their study designs and implementations multiple working hypotheses that might explain major reductions in breeding success (see also Nichols and Williams, 2006). And although we recognize in the multilevel mixed model a daunting challenge for the non-statistically inclined, we recommend that this model and other advancements provided through generalized linear models be embraced and promoted by the ecological research and teaching community as important alternative methods of data analysis (details are available in Singer and Willet, 2003) to time-honored approaches including analysis of variance and general linear models.

Acknowledgements

This research was sponsored in part by NOAA/NMFS (NA16FX1270) to CLB. We thank L. Baroff, SD, J. Brewer, T. Cooper, R. Fridinger, J.B. Gamble, D. McIntosh, K. M. Murra, R. Orben, N. Rojek, S. Studebaker, M. Van Sooy, S. Van Sooy and C.T. Williams for assistance in the field. We thank T. Drevon for donating time and expertise in developing our Visual Basic programs, K. Allard, J. Brewer, S.E. Whidden and C.T. Williams for valuable insights and discussions related to the preparation of this manuscript, and J.D. Singer for helpful suggestions.
REFERENCES


