Optimal sampling regime for detecting significant differences in peak mass of chicks: a case study with the wandering albatross *Diomedea exulans*

Steven G. Candy*, John van den Hoff

Australian Antarctic Division, 203 Channel Highway, Kingston 7050, Tasmania, Australia

ABSTRACT: Special emphasis should be placed on determining optimal sampling regimes to detect statistically significant differences in key indicators between populations for species with high conservation values, especially when invasive handling techniques are the only method of gathering data. In procellariiform nestlings one such key indicator is peak mass, which serves as a measure of parental investment and is an important factor for future survival. Daily measurements of mass for 15 wandering albatross *Diomedea exulans* chicks ranging in age from 14 to 270 d were obtained in 1963 on Bird Island, South Georgia. Log-transformed mass was modelled using a linear mixed model (LMM) that included a linear term for age combined with smooth nonlinear departures estimated as random effects, random chick effects and a continuous-time autoregressive error process. The relative efficiency of different sample sizes of chicks versus frequency of measurement was investigated by fitting this LMM to sub-samples of the dataset of either 5, 10 or 15 chicks with the corresponding frequency manipulated to give the same total number of measurements for each sub-sample. Measures of precision of predictions from the LMM obtained for each dataset included (1) the width of single standard error bounds about predicted peak mass as a percentage of this mass (PSEPM), and (2) the probability of detecting a *P*-percentage difference in peak mass between 2 theoretical populations for the same sample size of chicks and same frequency of measurement in each (PRDPM). The most efficient of these sampling regimes for detecting a difference between populations or cohorts was 15 birds sampled every third day, which gave a PRDPM of 0.95 for a *P* of 15%.

KEY WORDS: Optimal sampling · Albatross · Chick growth · Cubic splines · Autocorrelated data

INTRODUCTION

Information gathered from quantifying trends in aspects of the breeding success of procellariiform seabirds is important for managing the effects of incidental fishing mortality (Poncet et al. 2006). In addition, population drivers such as climate variability (Sydeman et al. 2001, Thompson & Grosbois 2002) and disease (Weimerskirch 2004) are important confounding effects on seabird populations, particularly as the former may alter ocean productivity. For birds, the maximum mass (i.e. peak mass) attained by chicks before fledging is a reflection of parental investment (or ecosystem productivity and variation) and potential

offspring survivorship. Because peak weight corresponds to a time when weight no longer increases it provides an indication of mature weight or maximal theoretical weight for a given season/year (Maruyama et al. 1999). The wide range of observed growth rates, including mass gain, among individual fork-tailed storm petrels *Oceanodroma furcate* was regarded as a response to environmental variability (Boersma & Parrish 1998), and environmental variability has been implicated in yellow-nosed albatross *Thalassarche chlororhynchos* chick provisioning parameters (Weimerskirch et al. 2001). Other species such as the wandering albatross *Diomedea exulans* are likely to respond in a similar fashion because breeding adult

birds are known to respond to changes in quality, distribution and abundance of their prey by producing changes in offspring tissue composition (e.g. weight) (Montevecchi 1993).

The wandering albatross has high conservation value (Croxall & Gales 1997) and as such, the ability to monitor changes in the peak average mass for chicks at particular colonies would provide information on regional population drivers, perhaps enabling distinctions between anthropogenic and natural forces to be better understood. Such monitoring requires that a sample of chicks be weighed throughout the season by manual handling since automatic nest weighing devices (e.g. Mulder & Swaan 1992, Huin et al. 2000) cannot be used. This is because wandering albatross chicks have a tendency to wander about the general nest area rather than remain on the nest for the full developmental period. Such behaviours reduce the effectiveness of automated measuring systems, and studies must then rely on more invasive handling techniques. Therefore, it is important to limit the sampling regime to the minimum number of samples needed to achieve the purpose of any study of growth, especially for species with high conservation values.

Estimating peak average mass for a colony and determining the precision of this estimate is important if this parameter is to be used repeatedly in a monitoring program. The precision of the estimate, and the associated ability to detect changes over space (e.g. different colonies, regions) and time (annual, decadal), is dependent on the sample size of chicks and the frequency of measurement. How these affect the efficiency of the sampling regime is, therefore, important in designing a cost-effective monitoring program (efficiency in this context is the degree of precision for a fixed sampling effort). Fitting smooth curves to measurements of mass for given times and finding the maximum mass from the fitted smooth curve, either analytically if it is a fully parametric model, such as the reverse Gompertz (Huin & Prince 2000), or by prediction at daily intervals from a semi-parametric model, such as a cubic smoothing spline, allows the estimate of peak average mass to 'gain strength' from, and 'smooth out' random errors in, the measurements around the time of the peak.

To investigate a method of estimating peak average mass and its precision as a function of sample size and frequency of measurement, daily measurements of mass for 15 wandering albatross chicks, gathered in 1963 on Bird Island, South Georgia, in the southern Atlantic Ocean (Tickell 1968) and ranging in age from 14 to 270 d, were modelled using a linear mixed model (LMM). The model incorporates a cubic smoothing spline in age constructed as the sum of a linear term in age and smoothed nonlinear departures estimated as random effects (Verbyla et al. 1999) to model nonlinearity in the average trajectory of logarithm of mass as a function of age. This LMM was used to estimate peak mass and other growth statistics and, via a sampling experiment, was used to investigate the relative efficiency of sample size of chicks versus frequency of measurement.

An alternative to the approach adopted here of fitting an average smooth curve across all chicks for mass versus age and finding the maximum from the fitted curve is the approach used by Huin & Prince (2000). They used a 2-stage random-coefficients approach, whereby in the first stage the reverse Gompertz model was fitted separately to each chick's data and quantities that are functions of the parameter estimates (i.e. peak mass) or predictions (i.e. growth rate, growth loss) obtained for each chick. Then, in the second stage these chick-level estimates were each analysed by means of univariate 2-way ANOVA accounting for species and year factors. However, this approach has limitations. Firstly, when the model cannot be successfully fitted for individual chicks (i.e. all model parameters cannot be simultaneously estimated successfully), the data for these chicks must be excluded. In addition, the second-stage univariate analysis does not correctly incorporate parameter estimation error at the first stage and ignores the between-parameter covariance for both the unobserved random parameters and their chick-specific estimates (Davidian & Giltinan 2003). A mixed model approach avoids these problems by combining random chick effects with fixed effects (such as island, year and species) in a single stage of estimation. In addition, cubic smoothing splines potentially provide more flexibility in modelling growth trajectories, especially when multiple feeding–fasting cycles occur.

As well as estimating various growth parameters, such as peak average mass, the LMM was used to investigate the relative efficiency of sample size of chicks versus frequency of measurement, by fitting it to sub-samples of the dataset using 3 sampling regimes (S) corresponding to combinations of number of chicks and interval between measurements. Measures of the precision of predictions from the fitted LMM for each dataset investigated were (1) the standard error (SE) of relative growth (SERG) between 10 and 180 d, (2) the area (AREA) between single SE bounds about predicted log mass over the 10 to 270 d range, (3) the width of single SE bounds about predicted peak mass as a percentage of this mass (PSEPM) and (4) the probability of detecting a P-percentage ($P\%$) difference in peak mass between 2 theoretical populations for the same sample size of chicks and frequency of measurement in each (PRDPM).

MATERIALS AND METHODS

Study area. The Bird Island study site, data collection methods and data structures are described in Tickell (1968). The growth trajectories for each of the 15 chicks measured are shown in Fig. 1. For ages in the ranges of 1 to 13 d and 271 to 302 d from hatch, the number of chicks measured ranged from 1 to 9, while for the 14 to 270 d range a minimum of 10 chicks were measured; for 204 of those daily ages at least 14 chicks were measured. Therefore, for the purpose of this study, since a close to balanced dataset was desirable, the data set used to fit the LMM and carry out the sampling efficiency study was restricted to birds with ages between 14 and 270 d, inclusive; this amounted to 3784 actual measurements. This dataset was 'padded out' for any remaining missing values of mass (i.e. insertion of a missing value for mass when not measured for a chick for any of the days in the above range) to give a balanced dataset of 3855 measurements (i.e. 15×257 daily observations, with mass values missing for 71 of them). This is the complete dataset used here. The data were 'padded out' in this way for 2 reasons. Firstly, this allowed the

variance structure in the LMM to be constructed efficiently using a direct product matrix formulation (Gilmour et al. 1999); this approach is well suited to such data where only a minor degree of imbalance exists. Secondly, it allowed a consistent definition, across chicks, of the time difference between selected measurements for the different frequencies of measurement (see next section).

Sampling regimes. Three sub-sampling regimes were investigated with the requirement that each would provide close to the same total number of measurements across chicks. The first regime used all 15 chicks but only retained every third daily measurement, giving an interval of 3 d between the measurements used in the analysis. This resulted in a single sub-sample (i.e. no randomisation) denoted as the S(15, 3 d) regime. The LMM was fitted, and precision statistics were calculated for this single sub-sample. The second regime, S(10, 2 d), sub-sampled 10 of the 15 birds (without replacement) and retained every second daily measurement for the 10 chicks. This regime was carried out 100 times with different random sub-samples of the 15 chicks for each replicate sample. The LMM was fitted, and precision statistics were calculated for each replicate dataset. The final regime, S(5, 1 d), was carried out as for the S(10, 2 d) regime, except that 5 of the 15 birds were sampled without replacement and each daily measurement for the 5 chicks was used.

Linear mixed model. The repeated observations of mass (i.e. mass versus age profiles) were modelled with LMMs incorporating cubic smoothing splines (Verbyla et al. 1999) fitted using the asreml library (Gilmour et al. 1995, 1999) within R (R Development Core Team 2008). The splines were fitted by asreml as the sum of a fixed-effect linear component plus a random-effect nonlinear component using every 5th day from Day 16 to Day 266 as a 'knot point' (Verbyla et al. 1999). The random term in the LMMs (apart from spline terms) was BIRD_ID with associated variance (on the log scale, see below) of σ_b^2. In addition, the fitted LMM incorporated the estimation of the autocorrelation between repeated measurements within chicks. The autocorrelation between observed ages within a chick's growth profile was modelled using an exponential power model with parameter φ so that the correlation between residuals at
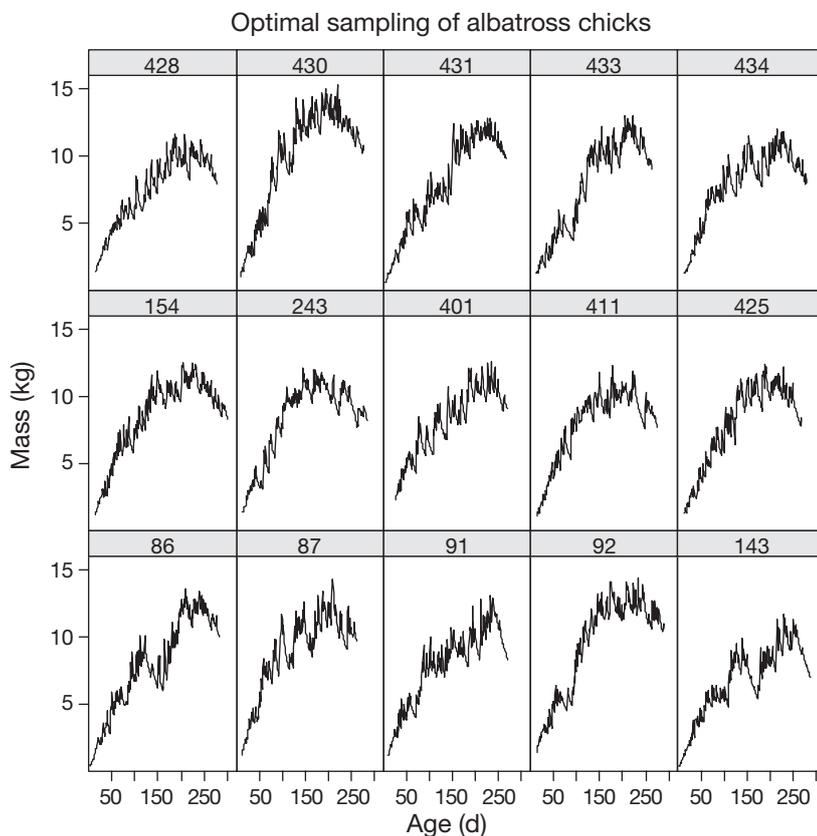

Fig. 1. *Diomedea exulans*. Individual growth trajectories for the 15 wandering albatross chicks weighed daily at Bird Island in 1963. Panel numbers denote the chick identity number

ages t and t' within a chick is given by $\varphi^{|t-t'|}$ (Gilmour et al. 1995, 1999) and, therefore, correctly handled missing values of mass (i.e. continuous-time exponential decay correlation equivalent for unit changes in time to a first-order autoregressive error model). This model corresponds to that of Diggle et al. (1994, p. 79) with experimental units (i.e. chicks) as random effects plus residual variance with autocorrelation but no measurement error.

To account for increasing variance of mass with time, mass data were log transformed so that the response variable fitted by the LMM was $y = \log(\text{Mass})$ and predictions on this scale, $\hat{y}$, could be back-transformed to give a predicted mass of $\exp(\hat{y})$. Approximate SEs of predicted mass were obtained as $\text{SE}(\hat{y})\exp(\hat{y})$ where $\text{SE}(\hat{y})$ is the SE on the transformed scale. The SE of predictions for each age (i.e. knot point) was calculated from the fitted LMM using the 'predict' function in the asreml library (for 'predict' methods see Welham et al. 2004) where the knot points in 'predict' can be changed from those used in the fit to be daily.

Measures of the precision of predictions. Once the LMM was fitted to a given dataset, measures of precision of estimates of various attributes of growth in mass were determined. These measures of precision and the corresponding attributes of growth are given below.

SERG: SERG is defined here as the SE of predicted relative growth between 10 and 180 d. This period was chosen because it is the approximate period between which mass increases from 10 to 90% of peak mass. Relative growth, R, for this period is given by:

$$R = \log_e\{M(t_{90})\} - \log_e\{M(t_{10})\} \tag{1}$$

where $t_{10} = 10$, $t_{90} = 180$, and $M(t)$ is mass (kg) at age t (d). The SE of R (SERG) was obtained simply as the SE of the difference between predicted $\log_e\{M(t)\}$ at these 2 ages obtained from the fitted LMM. The value of SERG is of practical interest because it determines, conditional on $M(t_{10})$, the SE of growth, G (g d^{-1}), over this period since

$$\text{SE}\{G|M(t_{10})\} \cong \text{SERG} \times 1000 \times M(t_{90})/(t_{90}-t_{10}) \tag{2}$$

where

$$G = 1000 \times \{M(t_{90})-M(t_{10})\}/(t_{90}-t_{10}) \tag{3}$$

This can be shown by noting that

$$G = 1000 \times M(t_{10})\{\exp(R)-1\}/(t_{90}-t_{10}) \tag{4}$$

and conditional on $M(t_{10})$ then

$$\text{var}\big(G|M(t_{10})\big) = \text{var}\{\exp(R)\}\big[1000 \times M(t_{10})/(t_{90}-t_{10})\big]^2 \tag{5}$$

and using a first order Taylor series approximation gives

$$\text{var}\{\exp(R)\} \cong \{\text{SERG} \times M(t_{90})/M(t_{10})\} \tag{6}$$

Combining these last 2 formulae and taking the square root gives the above formula for the SE of G.

AREA: To assess the precision of predictions of average mass over the entire age range the precision measure AREA was defined as the area between single SE bounds about predicted log mass over the range from 10 to 270 d. The width of single SE bounds was obtained for each day in the age range, $t = 10, 11, \ldots 270$, as $2 \times \text{SE}[\log_e\{M(t)\}]$, and AREA was calculated as the sum of these values.

PSEPM: A key quantity of interest is the peak average mass that was obtained by predicting log mass using the LMM and daily time steps between 10 and 270 d as above and finding the peak mass, $M(t_{100})$, and the age at which this occurred, t_{100}. PSEPM is defined here as the width of single SE bounds about predicted peak mass as a percentage of this mass and is given by:

$$100 \times \{\exp(\text{SE}[\log_e\{M(t_{100})\}])\} \tag{7}$$

PRDPM: Of interest to researchers is the ability to detect practically significant differences in peak mass between different populations and/or breeding years with a given degree of confidence (i.e. the power to detect the alternative hypothesis of a difference of given magnitude). The final precision measure, PRDPM, is defined here as the probability of detecting a $P\%$ difference in peak mass between 2 theoretical populations for the same sample size of chicks in each population (i.e. a total of twice the number of chicks in the dataset used in fitting the LMM) and the same frequency of measurement. Since the denominator degrees of freedom associated with test statistics obtained from predictions from complex LMMs are difficult to determine, the normal Z-statistic was used. Therefore, PRDPM was defined for the log of the ratio of the estimated peak mass for 2 populations (a,b) given by $\log_e\{M_a(t_{100})/M_b(t_{100})\}$ using the following probabilities

$$p_1(P,\alpha) = 1-\Phi\big(q_{1-\alpha}, \log_e\{1+P/100\}, \sqrt{2}\,\text{SE}\{\log_e(M(t_{100}))\}\big) \tag{8}$$

for alternative hypothesis H_1: $M_a(t_{100}) > M_b(t_{100})$, and

$$p_2(P,\alpha) = \Phi\big(q_\alpha, \log_e\{1-P/100\}, \sqrt{2}\,\text{SE}\{\log_e(M(t_{100}))\}\big) \tag{9}$$

for alternative hypothesis H_1: $M_a(t_{100}) < M_b(t_{100})$ where $\Phi(q_\alpha, u, s)$ is the cumulative probability density function for a normal distribution with mean u and standard deviation (SD), s, and $q_\alpha = \Phi^{-1}(\alpha, 0, s)$ for probability level $100(1-\alpha)$. If the ratio can be validly defined with either population denoted as a or b in the above ratio then PRDPM is best defined as the average $\{p_1(P, \alpha) + p_2(P, \alpha)\}/2$. For the present study this 'one-sided' power calculation was used, α was set to 0.05, and 2 values of P were investigated: 10 and 15%.

RESULTS

The average growth trajectory obtained from back-transformation of predictions from the fitted LMM with SE bounds obtained by applying the scalings $\exp(\pm SE[\log_e\{M(t)\}])$ to the average trajectory to give approximate upper and lower single SE bounds is shown in Fig. 2. The estimate of peak average mass, $M(t_{100})$, was 11.09 kg with approximate 95% confidence limits of 10.76 to 11.42 kg, and this peak was predicted to occur at Age 225 d.

The estimates for precision measures for the 3 sampling regimes and the complete dataset of daily measurements for all 15 chicks weighed at Bird Island are summarised in Table 1, which also gives the estimates, and for the S(10, 2 d) and S(5, 1d) regimes, the mean of estimates over replicates of the LMM parameters σ_b^2 and φ. The SD of these estimates obtained from the fit of the LMM to individual datasets is also given in Table 1. Again for the S(10, 2 d) and S(5, 1d) regimes these are means of estimates over replicates. The number of measurements (including missing values) for each of the 3 sampling regimes compared was close to 1285 in all replicates, which is one-third of the size of complete dataset of daily measurements for all 15 chicks.

Table 1. *Diomedea exulans*. Comparisons of the precision of predictions and variance parameter estimates from the linear mixed model (LMM) fitted to datasets derived from different sampling regimes for wandering albatross chicks weighed daily at Bird Island, South Georgia. The SD in parentheses determined for each statistic or element follows that value in the table (na: not available due to lack of replication). SERG: SE of predicted relative growth; AREA: area between single SE bounds about predicted log mass; PSEPM: width of single SE bounds about predicted peak mass; PRDPM: probability of detecting a $P\%$ difference in peak mass; $\hat{\sigma}_b^2, \hat{\varphi}$: LMM parameters

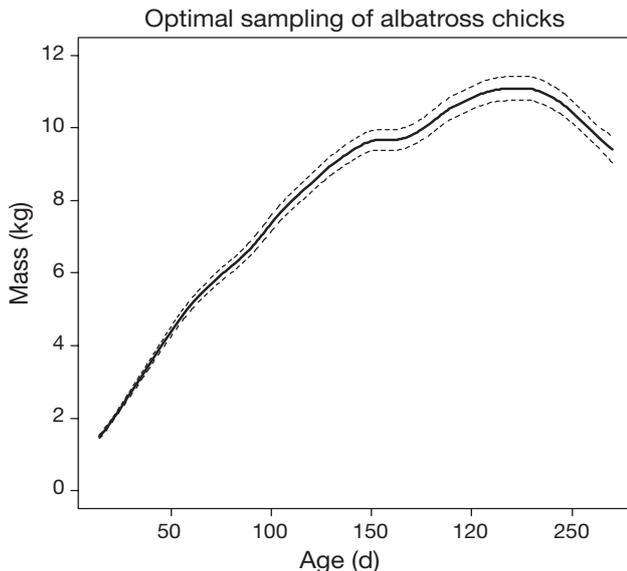| Statistic/estimate	Sampling regime No. of chicks (measurement interval, d)			
	15 (1)	15 (3)	10 (2)	5 (1)
SERG	0.0500	0.0529	0.0600	0.0730
	na	na	(0.0023)	(0.0057)
AREA	15.71	17.00	19.08	24.15
	na	na	(1.85)	(4.80)
PSEPM	5.91	6.41	7.18	9.05
	na	na	(0.72)	(1.86)
PRDPM ($P = 10\%$)	0.774	0.714	0.623	0.455
	na	na	(0.086)	(0.132)
PRDPM ($P = 15\%$)	0.972	0.949	0.888	0.748
	na	na	(0.045)	(0.140)
$\hat{\sigma}_b^2$	0.00672	0.00688	0.00847	0.01076
SE($\hat{\sigma}_b^2$)	0.00276	0.00294	0.00437	0.00748
$\hat{\varphi}$	0.776	0.844	0.820	0.744
SE($\hat{\varphi}$)	0.009	0.010	0.012	0.018

Fig. 2. *Diomedea exulans*. Average growth trajectory for wandering albatross chicks at Bird Island obtained from back-transformation of predictions from the fitted linear mixed model showing single SE bounds (fine dashed lines)

DISCUSSION

The results in Table 1 show that for all 5 measures of precision the 15 chick sample size gave the best precision. If the number of chicks is fixed (e.g. 15) then the comparison of the results for the S(15, 3d) regime to the complete dataset of 15 chicks and daily measurements shows that daily measurements do improve precision most notably for measures AREA, PSEPM and PRDPM. This is worth noting, especially for PSEPM, which is basically the precision of an estimate of mass at a single age and, therefore, more frequent measurements might seem less important for improving precision. However, the cubic spline components of the LMM give predictions that 'gain strength' from considering the growth profile as a sequence of values that follow a clear trend, rather than simply a set of daily means. There was an obvious increasing trend in values of SE($\hat{\sigma}_b^2$) and SE($\hat{\varphi}$) as sample size of chicks was reduced from 15 to 5 (Table 1). Similarly, there was a strong decreasing trend in PRDPM as number of chicks was reduced with the mean value for S(5, 1 d) scenario for P of 10% as low as 0.455. There was a similar reduction in power for a 15% difference.

This study of sampling efficiency used sub-samples of actual data to model the relationship between precision of estimates and sampling regime. Alternatively, Monte Carlo sampling could have been used by drawing samples from normal distributions for log values of mass both within and between chicks using the fitted

LMM parameters for the complete dataset to generate these distributions. In this way, sample sizes of chicks larger than 15 could be investigated. However, the advantage of the sub-sampling approach adopted here is that inferences were drawn from real data even if the sampling regimes investigated were more limited than those possible with the Monte Carlo method. The limitation of the Monte Carlo method is that the real data includes stochastic processes not fully captured by the LMM, such as asynchronous satiation and fasting periods across chicks, which result in sudden increases and decreases in mass, respectively. Combined with this is the complexity introduced by different hatching times of chicks (i.e. day of the year and age do not exhibit a one-to-one correspondence across chicks), giving rise to complex patterns in within-chick residuals from the LMM. These trends were modelled to some degree by the flexibility of cubic smoothing splines combined with the estimation of chick random effects, and the autoregressive within-chick error structure, but clearly not all of the above complexity could be expected to be subsumed by the LMM. For this reason an empirical study using sub-sampling of real data was preferred to a Monte Carlo approach.

Studies of sampling efficiency in growth studies, such as that described here, need to adequately model the error structure of the data for the efficiency of alternative sampling strategies to be accurately assessed. The incorporation of the autoregressive error structure for the within-chick residuals using the LMM is important for correctly quantifying the information content of daily versus less frequent measurements.

Clearly, if it is necessary to make a trade-off between the number of chicks measured compared with frequency of measurement, given that the total number of measurements is limited by available sampling time when manually weighing chicks, then measuring more chicks less frequently is a more effective strategy. However, this study has not investigated a sampling frequency of less than every third day. Very infrequent measurement can reduce precision significantly, as seen for example in the precision for wandering albatross chicks weighed at Macquarie Island (authors' unpubl. data). Therefore, neither few measurements per chick nor a small sample size of chicks are efficient sampling strategies and, where possible, such sampling strategies should be avoided as the data collected do not warrant the potential disturbance to the birds. Given the results of this empirical study, measurement every third day of at least 15 chicks per cohort or population is recommended for growth studies of wandering albatross chicks. If more resources are available than this, given the trends discussed (Table 1), these should be invested in measuring more chicks (where possible) rather then increasing the

frequency of measurement. Growth studies of other species could benefit from similar studies of sampling efficiency.

In studies that were able to use automatic nest weighing devices (e.g. Mulder & Swaan 1992, Huin et al. 2000), the cost of more frequent measurements becomes zero, and the main consideration becomes how to adjust weight measurements for spurious effects (e.g. due to the movement of the chick, wet versus dry chick or the effect of wind) and how often calibration is required (i.e. due to losses or additions to the nest such as vomit and snow). However, many procellariiforms (e.g. petrels) either nest underground in burrows (e.g. shearwaters and prions), in rock crevices (e.g. snow petrels) or on bare rock (e.g. fulmars) where automatic nest weighing is less practical than manual restraint. If a trade-off between number of chicks and frequency of measurement is not directly relevant in studies using automatic nest-weighing devices, the sample size required to achieve a desired power for the test given by the measure PRDPM is important information for such studies.

LITERATURE CITED

Boersma PD, Parrish JK (1998) Flexible growth rates in fork-tailed storm-petrels: a response to environmental variability. Auk 115:67–75

Croxall JP, Gales R (1998) An assessment of the conservation status of albatrosses. In: Robertson G, Gales R (eds) Albatross biology and conservation. Surrey Beattie & Sons, Chipping Norton, p 46–65

Davidian M, Giltinan DM (2003) Nonlinear models for repeated measurement data: an overview and update. J Agric Biol Environ Stat 8:387–419

Diggle PJ, Liang K, Zeger SL (1994) Analysis of longitudinal data. Clarendon Press, London

Gilmour RA, Thompson R, Cullis BR (1995) Average information REML: an efficient algorithm for variance parameter estimation in linear mixed models. Biometrics 51: 1440–1450

Gilmour RA, Cullis BR, Welham SJ, Thompson R (1999) Asreml reference manual. Biometric Bulletin No. 3. New South Wales Agriculture, Orange Agricultural Institute, Orange

Huin N, Prince PA (2000) Chick growth in albatrosses: curve fitting with a twist. J Avian Biol 31:418–425

Huin N, Prince PA, Briggs DR (2000) Chick provisioning rates and growth in blackbrowed albatross *Diomedea melanophris* and grey-headed albatross *D. chrysostoma* at Bird Island, South Georgia. Ibis 142:550–565

Maruyama K, Akbar MK, Turk CM (1999) Growth patterns and carcass development in male ducks selected for growth rate. Br Poult Sci 40:233–239

Montevecchi WA (1993) Birds as indicators of change in marine prey stocks. In: Furness RW, Greenwood JJD (eds)

Birds as monitors of environmental change. Chapman & Hall, London, p 217–266

Mulder JL, Swaan AH (1992) Body-weight changes of egg-laying curlews *Numenius arquata*, as monitored by an automatic weighing system. Ardea 80:273–297

Poncet P, Robertson G, Phillips RA, Lawton K, Phalan B, Trathan PN, Croxall JP (2006) Status and distribution of wandering, black-browed and grey-headed albatrosses breeding at South Georgia. Polar Biol 29:772–781

R Development Core Team (2008). R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, www.R-project.org

Sydeman WJ, Hester MM, Thayer JA, Gress F, Martin P, Boffa J (2001) Climate change, reproductive performance and diet composition of marine birds in the southern California Current system, 1969–1997. Prog Oceanogr 49:309–329

Thompson PM, Grosbois V (2002) Effects of climate change on seabird population dynamics. Dir Sci 1:50–52

Tickell WLN (1968) The biology of the great albatrosses, *Diomedea exulans* and *Diomedea epomophora.* In: Austin OL (ed) Antarctic bird studies. Am Geophys Union Antarc Res Ser 12, Washington, DC, p 1–55

Verbyla AP, Cullis BR, Kenward MG, Welham SJ (1999) The analysis of designed experiments and longitudinal data using smoothing splines (with discussion). Appl Stat 48: 269–311

Weimerskirch H (2004) Diseases threaten Southern Ocean albatrosses. Polar Biol 27:374–379

Weimerskirch H, Zimmermann L, Prince PA (2001) Influence of environmental variability on breeding effort in a long-lived seabird, the yellow-nosed albatross. Behav Ecol 12: 22–30

Welham SJ, Cullis BR, Gogel B, Gilmour A, Thompson R (2004) Prediction in linear mixed models. Aust NZ J Stat 46:325–347