Factors influencing heterogeneity in female reproductive success in a Critically Endangered population of bottlenose dolphins

T. E. Brough1,*, S. Henderson1,2, M. Guerra1, S. M. Dawson1

1Marine Mammal Research Group, Department of Marine Science, University of Otago, PO Box 56, Dunedin, New Zealand
2Department of Biology, Tacoma Community College, 6501 South 19th Street, Tacoma, Washington 98466, USA

ABSTRACT: For threatened species or populations, variation in reproductive success among females may be explicitly linked with vulnerability to extinction. Thus, an understanding of factors that may cause variability in reproductive success is important. The population of bottlenose dolphins in Doubtful Sound, New Zealand, has a recent history of rapid population decline and low calf survival rates. A previous study has shown high variability in calf survival among multiparous females. This study addresses the factors that seem most important in explaining variation in calf survival and thus reproductive success among females in this population. Reproductive data were sourced from a long-term photo-identification dataset, which allowed tracking the fate of 49 calves born into the population between 1995 and 2012. General linear mixed models combined with model averaging were used to assess how birth timing, maternal size, age and potential anthropogenic impacts contributed to variation in calf survival. Models show that a female’s size and her ability to give birth at an optimum time in the calving season are significant predictors of calf survival to an age of 1 and 3 yr. This is the first study to demonstrate how birth timing and mother size are correlated with female reproductive success in a cetacean species. These results confirm the importance of demographic stochasticity and reproductive heterogeneity in small, threatened marine mammal populations.

KEY WORDS: Reproductive success · Reproductive heterogeneity · Survival · Bottlenose dolphin

INTRODUCTION

Understanding the factors influencing reproduction in threatened species may provide valuable insights into the causes of low population sizes and overall vulnerability to extinction (Craig & Ragen 1999, Baker et al. 2007, Berger 2012). In population biology, one of the critical parameters of interest is reproductive success, which is inherently associated with the productivity and thus viability of a population (Andrén 1990, Kendall & Fox 2002, Melbourne & Hastings 2008). Reproductive success is often defined as the number of offspring that recruit into the adult breeding population (McGregor et al. 1981, Phillips et al. 1996). In other cases, differences in reproductive success among females are described in terms of the number of offspring that survive to a particular age (Mann et al. 2000, Wells et al. 2005, Henderson et al. 2014).

Reproductive success may be influenced by a range of social (Côtè & Festa-Bianchet 2001), ecological (Atkinson & Ramsay 1995, Festa-Bianchet & Jørgenson 1998), behavioural (Mann et al. 2000) and morphological or physiological factors (Pomeroy et al. 1999). For example, over a wide range of taxa, offspring survival is correlated with maternal body size (Chastel et al. 1995, Fairbanks & McGuire 1995, Pomeroy et al. 1999). Large females, for example, probably have greater resources to invest in their young during embryonic and/or post-natal development (e.g. grey seals, Iverson et al. 1993; polar bears, Atkinson & Ramsay 1995). The timing of births may also be correlated with female reproductive success (Goldizen et al. 1988, Majluf 1992). It has been sug-
gested that this relationship is due to benevolent environmental conditions and/or resource availability increasing the likelihood of offspring survival at particular times (Ims 1990, Ransome & McOwat 1994, Daan et al. 1988).

Individuals in many species show variation in reproductive success according to age (e.g. sea otters, Riedman et al. 1994; polar bears, Atkinson & Ramsay 1995; southern right whales, Elwen & Best 2004). Such a relationship may represent older, more experienced females retaining a higher position in the social hierarchy and thus being exposed to better mating opportunities (Côté & Festa-Bianchet 2001). Additionally, more experienced females may ‘learn’ the optimal conditions for giving birth through adaptive trial and error (Clutton-Brock 1984).

The effects of maternal size, birth timing and maternal age on reproductive success are largely unknown in cetaceans. To our knowledge, there has been no previous investigation of whether larger females exhibit greater reproductive success. This probably reflects the difficulty of gathering the required morphometric data. Cetaceans are among the most difficult animals to measure accurately in the wild (Dawson et al. 1995). Whilst strong birth seasonality is evident in many cetacean populations (Mann et al. 2000, Thayer et al. 2003) and studies have suggested a relationship between reproductive success and birth timing (Henderson et al. 2014, Fruet et al. 2015), the link is yet to be demonstrated.

Bottlenose dolphins are a species which is widely distributed globally and is the subject of several long-term (>20 yr) research programmes (e.g. Scott et al. 1990, Wilson et al. 1999, Mann et al. 2000, Henderson et al. 2014). Knowledge of their reproductive biology in the wild, however, is limited to documentation of reproductive rates (Mann et al. 2000, Henderson et al. 2014), birth seasonality (Urián et al. 1996, Thayer et al. 2003) and interbirth intervals (Mann et al. 2000, Henderson et al. 2014, Fruet et al. 2015). Reproduction is likely to be influenced by a range of factors that result in variation in reproductive success among years and individuals. For example, Mann et al. (2000) demonstrated that female reproductive success was influenced by the use of shallow water habitat, which may be associated with reduced calf predation or increased food resources.

The bottlenose dolphin population of Doubtful Sound has been the subject of ongoing research since 1990 (Williams et al. 1993) and is currently classified as Critically Endangered (Currey et al. 2009a). The long-term data set generated from decades of photo-identification mark-recapture sampling provides the opportunity to produce detailed reproductive histories for each of the females in the population. The population has suffered periods of dramatic population decline (Currey et al. 2007), the demographic cause of which has been attributed to a decrease in calf survival rate from 0.86 (95% CI: 0.68−0.95) prior to 2002 to 0.38 (95% CI: 0.21−0.58) for the period 2002 to 2007 (Currey et al. 2009b). Potential anthropogenic impacts due to habitat modification from a hydro-electric power station have been discussed (Currey et al. 2009a,b).

Henderson et al. (2014) demonstrated substantial heterogeneity in reproductive success among female dolphins in Doubtful Sound. Of the 18 multiparous females in the population at 2011, 6 had not had a calf survive to 3 yr of age since 1995. In contrast, another 6 females had 70% of their calves survive between 1995 and 2011. This is despite similar numbers of calves being born to both ‘good’ and ‘bad’ mothers (Henderson et al. 2014). Calf survival in Doubtful Sound seems to be higher if a calf is born in January, which is immediately before the peak in surface water temperature (Henderson et al. 2014). This population resides near the southernmost limit for the species and is thus faced with much cooler temperatures than bottlenose dolphins in lower latitudes (Haase & Schneider 2001). Dolphin calves are less tolerant of low water temperatures than adults due to their proportionally greater surface area and reduced blubber thickness (Yeates & Houser 2008). The thermal stresses imposed by being born into a cool water environment may cause high calf mortality for individuals not born at the optimum time (Henderson et al. 2014).

This study assessed how a range of measurable factors (mother size, mother age, and birth timing) and potential anthropogenic impacts contribute to female reproductive success. This was undertaken by assessing calf survival to 1 and 3 yr of age for each of the reproducitively active females in the population of Doubtful Sound. Quantifying how these factors influence female reproductive success will facilitate understanding the drivers of population trends in Doubtful Sound, which are likely to be relevant to other small, threatened marine mammal populations.

**MATERIALS AND METHODS**

**Photo-identification surveys**

Photo-identification surveys of bottlenose dolphins in Doubtful Sound (Fig. 1) have been carried out...
since 1990 (Williams et al. 1993) following a predefined route. Photographs were taken using Nikon SLR and DSLR cameras, mostly with 80–200 mm f2.8 and 300 mm f4 AF Nikkor lenses. Field and photo-identification protocols are described in detail in Williams et al. (1993) and Currey et al. (2009b).

Reproductive data

Reproductive histories over the period between 1995 and 2012 were available for all adult females in the Doubtful Sound population that were still alive at the end of 2012 (n = 19). These females gave birth to 49 calves that were included in this analysis. Due to a high frequency of sampling trips (mean = 14 mo⁻¹) over the months during which calves are born (September to April; see Supplement 1 at www.int-res.com/articles/suppl/n029p255_supp.pdf for research effort), birth month is known for most calves, and we can assume that a calf is sighted during its month of birth. Calves in Doubtful Sound are almost never born outside this 8 mo period (Haase & Schneider 2001, Henderson et al. 2014). Although there has been some extension of home-range beyond Doubtful Sound over recent years, the population is essentially closed to immigration/emigration and all individuals have very high resighting rates, with every individual in the population typically sighted each trip (Henderson et al. 2013). Consequently, we assumed that if a female was repeatedly resighted (minimum of 10 encounters) without her calf during a monitoring trip after the calf had been first sighted, the calf had died. The identity of a calf’s mother was inferred from constant close association during a minimum of 10 separate encounters over several days, of a particular female with that calf. The small size of the population (n = 60 at 2012) and high resighting rates allow for confidence that every reproductively active female (provided they are over 8 yr of age [see ‘Mother size’ below]) has been included in the analysis. Because of the long-term nature of the study, in order for a female not to be included as reproductively active, the survey team would have had to have missed multiple calf births for that female. This is very unlikely. Moreover, the sex of every adult individual in the population is known via close observation over 20+ yr. Reproductive histories were used to assign a fate (i.e. survival to 1 and 3 yr), time of birth and mother ID to each calf.

Mother size

The size of female dolphins was measured in the field using a purpose-built digital stereo-photogrammetric system (Brough 2013), which is a modified version of that used to measure sperm whales (Grov-cott et al. 2012). The method allowed for repeated measurement of individuals to assess the degree of measurement error inherent within the system (mean CV = 2.43%). The measured distance (upper jaw to dorsal fin, UJ–DF) is a good proxy for total length (r² = 0.94, Brough 2013). Total length is a standard measurement of size in marine mammals which is closely related to total body mass (Trites & Pauly 1998). Animal condition is likely to change from year to year (Miller et al. 2011, 2012), and thus our static (at 2012) length estimate for female dolphins is unlikely to represent a mother’s breeding condition for calf births before 2012. While year-to-year breeding condition is an important factor influencing reproductive success (Miller et al. 2012), variation in absolute body size among individuals in a population
may also be important (Reiter et al. 1981, Hussey et al. 2010). In Doubtful Sound there is substantial variation in body size among adult dolphins (Chong & Schneider 2001, Brough 2013), providing an interesting opportunity to assess how reproductive success may be influenced by body size in a cetacean species inhabiting a harsh environment. In this population, the majority of body growth occurs during the first 8 yr of life (Chong & Schneider 2001). Thus, in order to ensure that the size of individual females has remained constant over the study period, calf births were excluded from the analysis if mothers could have been younger than 8 yr.

### Month of birth

The long-term dataset included the dates of first and last sightings for all calves born into the population (Henderson et al. 2014). Cases in which birth month was not certain (e.g. when there was >1 mo between sampling trips) were not included in the analysis. Preliminary analysis of month of birth data indicated problems with model convergence when using birth month (October–April) as different levels in a categorical factor. This is likely due to there being very few observations of calf births in months other than December, January and February and the subsequent problems with data dispersion. For this reason the month of birth factor clustered data into 3 levels describing whether calves were born before January (the months October, November and December), during January, or after January (the months February, March and April). No calf births events were recorded between May and September. This clustering of birth month data is consistent with a previous study on this population that suggests January may be a ‘sweet-spot’ for calf survival (Henderson et al. 2014).

### Mother age

Age is known directly for females born into the population since 1990. Ten of the reproductively active females in the population were, however, first observed as ‘adults’ (>3 yr old) at the beginning of the research programme and were therefore accorded a value for minimum age in the analysis (i.e. the age value for these females was set at 3 yr at 1990). Conceivably, these 10 females could have true ages anywhere between 3 and 50 yr, and so this decision obviously reduces our ability to assess whether the older females have reduced reproductive success. The age data distinguished between the older females and those born into the population, thus presenting an opportunity to assess if younger females are more or less successful at raising young. In some species, disparity in reproductive success due to age is most evident for young females, especially when full reproductive potential is not reached until well after age of first birth (Clutton-Brock 1984, Cameron et al. 2000). Thus, including an age parameter in this analysis that distinguishes young females was considered useful.

### Tailrace freshwater discharge

Doubtful Sound is subjected to habitat modification due to the existence of a hydro-electric power station that diverts up to 550 m$^3$ s$^{-1}$ of freshwater from Lake Manapouri into the fiord (Bowman et al. 1999). It may be that the increased freshwater discharge from the power station affects both temperature and resources in the fiord. Such an affect could be correlated with the significant decrease in calf survival at 2002, when a second tailrace tunnel was opened (Currey et al. 2009b). The second tailrace tunnel from the Lake Manapouri hydro-electric scheme began to divert water into the Doubtful Sound complex in May 2002. In order to assess whether calf survival has decreased following the opening of the second tailrace, a binomial predictor variable was included in the modelling structure that specifies whether the second tailrace was operational (0) or not (1).

### Statistical approach

This study used general linear mixed models (GLMM) to assess the range of effects contributing to variation in calf survival to 1 yr (‘S.1’) or to 3 yr (‘S.3’). In Doubtful Sound, calf mortality typically occurs within the first year (usually in the first month) with an additional peak in mortality around 3 yr of age, possibly associated with weaning (Henderson et al. 2014). Other studies assess reproductive success of mothers as raising a calf to 3 yr old only (Mann et al. 2000, Mann & Watson-Capps 2005, Mann et al. 2008). However, as the various factors may differ in importance as calves grow, we chose to model variation in survival to both 1 and 3 yr of age.

The GLMM framework allows for a combination of categorical, continuous and binomial data and enables the inclusion of both fixed and random effects.
input parameters were created for each global model, and are becoming increasing popular for modeling survival data with random (e.g. individual level) factors (Milner et al. 1999, Vergara et al. 2007).

The model framework was based on a binary response variable of calf survival (1 = survived, 0 = did not survive). There was an observation for every calf born into the population between 1995 and 2012, provided its mother was known, was still present in 2012, and the mother was over the age of 8 yr at the time of birth (see above). Calves whose mothers had died before 2012 (n mothers = 8) were not included in the study. As photogrammetric sampling was carried out in 2012, size estimates were available only for females that were alive in 2012.

Each observation of a calf survival event had associated observations of various factors. Interactions effects between parameters were excluded from the modelling framework due to issues with model convergence that were potentially a product of generating complex models with a dataset that was (unavoidably) small (Zuur et al. 2009). Mother ID was specified as a random factor (‘RE’) in the modelling framework.

Models were developed in the program R (version 3.0.3; R Core Team 2015) using the package lme4 (version 1.1-5; Bates et al. 2014a) following a method described by Grueber et al. (2011). The glmer function of lme4 fits GLMM models by estimating model parameters via maximum likelihood (ML). ML is approximated using either Laplace quadrature or Gaussian-Hermite quadrature (Bolker et al. 2009a, Zuur et al. 2009). Two global models were configured, one for each response variable (S.1 and S.3), that included each input parameter as below:

\[ S.1 \sim \text{Size} + \text{Age} + \text{Month}(\text{factor}) + \text{Tailrace} \]

\[ S.3 \sim \text{Size} + \text{Age} + \text{Month}(\text{factor}) + \text{Tailrace} \]

Mother size (Size) and mother age (Age) are discrete continuous variables, Month is a categorical factor with 3 levels and Tailrace is binomial. The models used a logit link function, binomial error distribution and Laplace approximation of ML. Model data were standardised using the centralising mean method with the standardize function in package arm (version 1.6-10; Gelman & Su 2013). Next, full model sets containing every possible combination of input parameters were created for each global model using the dredge function in MuMIn (version 1.9.0; Barto 2015). We used an information-theoretic approach to model selection (Burnham & Anderson 2002). Models were ranked via AICc, a modified version of Akaike’s information criterion (Akaike 1973) which reduces bias for small sample sizes (Hurvich & Tsai 1989). A ‘top model’ set was generated for each response that included every model having a ΔAICc value of 6 or less compared to the most parsimonious model (Burnham & Anderson 2002).

In cases in which differences among subjects are not of particular interest, random factors are often treated as ‘nuisance’ parameters (Bolker et al. 2009b, Iwasaki & Brinkman 2015). In the context of this study, differences in reproductive success among female dolphins are of interest. Allowing for a RE parameter of mother ID in our modelling framework presents an opportunity to assess this. To assess the level of variation in reproductive success among female dolphins, we compared the model outputs and model selection values (AICc values and model weights) of a reduced (null) model containing only our RE parameter, with the top models for both S.1 and S.3 model sets. RE (among-subject) variation can be accounted for by variation in fixed effects (Scheipl et al. 2008, see ‘Results’), which may present difficulties when using models with significant fixed effects to assess differences among RE levels. Therefore, it was necessary to use the conditional modes and unconditional standard errors from the S.1 and S.3 null models to evaluate differences in reproductive success among individual females (see Table 3). The conditional modes and unconditional error of each level within our random factor were extracted using the ranef function in package lme4.

To assess the fit of the models to the survival data, diagnostic plots were generated to assess (1) homogeneity of variance, (2) whether data transformed by the link function were linear with respect to continuous input parameters and (3) normality of model residuals and random effects (Bolker et al. 2009a, Bolker 2015). The binomial response variable made it necessary to fit smoothed curves with confidence intervals to plots of residuals (Pearson) vs. fitted values to determine (1). This was undertaken with package qgplot2 (version 1.0.1; Wickham & Chang 2015) using the functions qplot and stat_smooth. The same function was used to assess (2) by plotting continuous input parameters (size & age) against the models’ predicted values. A further assessment of (1) was carried out using binned residual values vs. expected model values using binnedplot in the package arm. Normality of residuals (3) was assessed
using the function \texttt{qqnorm} to compare the distribution of model residuals with a normal distribution. The assumption that the random effects assume a normal distribution was also tested with \texttt{qqnorm} by comparing the conditional modes of the random parameter with a normal distribution. To further test residual normality Shapiro-Wilk tests (Shapiro & Wilk 1965) were carried out on both the model residuals and RE conditional modes. Data overdispersion was assessed for both model sets by comparing residual deviance with residual degrees of freedom (Bolker 2015).

**Model averaging**

For model outcomes with no clear 'best' model (model weight 0.90 or over and/or \( \Delta \text{AICc} < 2 \)), parameter estimates are more realistic when the coefficients are averaged across all models that contribute some weight, rather than relying on estimates generated from the top model only (Burnham & Anderson 2002, Johnson & Omland 2004).

Two main model-averaging methods have been advanced (Burnham & Anderson 2002). The ‘natural-average’ method averages parameters and error over models in which they appear, weighting each parameter by the sum of the model weights. In contrast, the zero method averages each parameter over all models in the set, substituting a 0 value when a parameter is absent from a particular model. (Burnham & Anderson 2002, Grueber et al. 2011). It has been suggested that the ‘zero’ method is more appropriate when the goal is to determine the factors that have the strongest influence on the response (Nakagawa & Freckleton 2011). The ‘natural average’ method is more suited to situations in which there is a particular variable of interest may have a comparatively small effect size; thus, the shrinkage encountered with the ‘zero’ method is probably best avoided (Nakagawa & Freckleton 2011).

Model averaging of parameters was carried out with the \texttt{model.avg} function in \texttt{MuMIn} using the ‘zero method’, because this study is concerned with determining the factors that have the strongest influence on the response (Burnham & Anderson 2002). The size of the estimated coefficient indicates the degree to which each predictor accounts for variation in calf survival and allows the predictor variables to be ranked. We used 95% confidence intervals to judge the significance of each parameter's predictive capacity (Burnham & Anderson 2002, Grueber et al. 2011, Nakagawa & Freckleton 2011). A parameter was deemed significant if its confidence interval did not include zero (Grueber et al. 2011, Nakagawa & Freckleton 2011).

The inclusion of models in a top model set that are more complex versions of a highly ranked model may cause bias when performing model averaging (Richards 2005, 2008, Richards et al. 2011). This problem can be exacerbated with binomial datasets and can be particularly problematic with overdispersed data (Richards 2005, 2008). Judging by traditional tests, there is no indication that the data in this study were overdispersed (see ‘Results’); however, we undertook further assessment of overdispersion because of the potential for it to have substantial effects on model averaging results when complex models are included in the set. Following a method suggested by Richards (2008), models that were more complex versions of higher ranked models in the set (i.e. models that were nested within a simpler model with smaller \( \Delta \text{AICc} \)) were removed from the ‘top model’ set (Richards 2008). Model averaging was then performed on the remaining models (Smodes) and the subsequent parameter estimates compared to those produced using the entire ‘top model’ set. Sizeable differences in parameter estimates, or in the statistical significance of input parameters, would suggest that model averaging across a top model set may not be appropriate given the dispersion of the data (Richards 2005, 2008). Such differences could be associated with including overly complex models in the ‘top model’ set.

**RESULTS**

**Observations**

As of 2012, there were 19 reproductively active females in the Doubtful Sound population. These females had given birth to a total of 57 calves since 1995. Birth month of all but 2 calves was known with confidence, and 6 of these calves were born when mothers could have been <8 yr of age. Thus, the total sample size for analysis was 49. Thirty-four of the 49 survived their first year, whilst 27 of 49 survived to 3 yr of age.

Photogrammetric estimates of body size were available for all mothers currently in the population. There was substantial variation in measurements of mother size with a minimum UJ-DF distance of 132 cm and a maximum of 217 cm (mean = 163 cm). Data on body proportions suggest that these measurements correspond to total lengths of 219 and
Table 1. Summary of the ‘top model’ set used to estimate parameters for calf survival to 1 yr. Fixed input parameters include mother size (Size), month of birth (Month), mother age (Age) and Tailrace. The null model (i.e. no fixed effects), containing the random effect of mother ID (Mother) was not included in the top model set, yet is shown in the table for comparison. This model selection table summarises the top models in terms of degrees of freedom (df), Akaike information criterion (AICc), delta AICc (ΔAICc), AICc model weight (AICc.Wt) and a cumulative weight index (Cum.Wt).

<table>
<thead>
<tr>
<th>Model</th>
<th>Rank</th>
<th>df</th>
<th>AICc</th>
<th>ΔAICc</th>
<th>AICc.Wt</th>
<th>Cum.Wt</th>
</tr>
</thead>
<tbody>
<tr>
<td>Month + Size</td>
<td>1</td>
<td>5</td>
<td>54.59</td>
<td>0</td>
<td>0.34</td>
<td>0.34</td>
</tr>
<tr>
<td>Size</td>
<td>2</td>
<td>3</td>
<td>55.65</td>
<td>1.06</td>
<td>0.20</td>
<td>0.53</td>
</tr>
<tr>
<td>Month + Tailrace + Size</td>
<td>3</td>
<td>6</td>
<td>56.37</td>
<td>1.99</td>
<td>0.12</td>
<td>0.66</td>
</tr>
<tr>
<td>Month + Age + Size</td>
<td>4</td>
<td>6</td>
<td>56.66</td>
<td>2.07</td>
<td>0.12</td>
<td>0.78</td>
</tr>
<tr>
<td>Tailrace + Size</td>
<td>5</td>
<td>4</td>
<td>57.69</td>
<td>3.11</td>
<td>0.07</td>
<td>0.85</td>
</tr>
<tr>
<td>Age + Size</td>
<td>6</td>
<td>4</td>
<td>57.79</td>
<td>3.20</td>
<td>0.07</td>
<td>0.92</td>
</tr>
<tr>
<td>Month + Tailrace + Age + Size</td>
<td>7</td>
<td>7</td>
<td>59.26</td>
<td>4.68</td>
<td>0.03</td>
<td>0.95</td>
</tr>
<tr>
<td>Month + Tailrace + Age + Size</td>
<td>8</td>
<td>7</td>
<td>59.48</td>
<td>4.90</td>
<td>0.03</td>
<td>0.98</td>
</tr>
<tr>
<td>Tailrace + Age + Size</td>
<td>9</td>
<td>5</td>
<td>60.17</td>
<td>5.59</td>
<td>0.02</td>
<td>1.00</td>
</tr>
<tr>
<td>Mother (null)</td>
<td>12</td>
<td>2</td>
<td>61.63</td>
<td>7.05</td>
<td>0.01</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Table 2. Summary of the ‘top model’ set used to estimate parameters for calf survival to 3 yr. Input parameters and abbreviations as in Table 1.

<table>
<thead>
<tr>
<th>Model</th>
<th>Rank</th>
<th>df</th>
<th>AICc</th>
<th>ΔAICc</th>
<th>AICc.Wt</th>
<th>Cum.Wt</th>
</tr>
</thead>
<tbody>
<tr>
<td>Month + Size</td>
<td>1</td>
<td>5</td>
<td>55.9</td>
<td>0</td>
<td>0.27</td>
<td>0.27</td>
</tr>
<tr>
<td>Size</td>
<td>2</td>
<td>3</td>
<td>56.19</td>
<td>0.29</td>
<td>0.23</td>
<td>0.50</td>
</tr>
<tr>
<td>Tailrace + Month + Size</td>
<td>3</td>
<td>6</td>
<td>57.16</td>
<td>1.25</td>
<td>0.14</td>
<td>0.64</td>
</tr>
<tr>
<td>Month + Age + Size</td>
<td>4</td>
<td>6</td>
<td>57.69</td>
<td>1.79</td>
<td>0.11</td>
<td>0.75</td>
</tr>
<tr>
<td>Tailrace + Size</td>
<td>5</td>
<td>4</td>
<td>58.02</td>
<td>2.12</td>
<td>0.09</td>
<td>0.85</td>
</tr>
<tr>
<td>Age + Size</td>
<td>6</td>
<td>4</td>
<td>58.11</td>
<td>2.21</td>
<td>0.09</td>
<td>0.94</td>
</tr>
<tr>
<td>Tailrace + Month + Age + Size</td>
<td>7</td>
<td>7</td>
<td>59.90</td>
<td>3.99</td>
<td>0.04</td>
<td>0.97</td>
</tr>
<tr>
<td>Tailrace + Age + Size</td>
<td>8</td>
<td>5</td>
<td>60.48</td>
<td>4.58</td>
<td>0.03</td>
<td>1.00</td>
</tr>
<tr>
<td>Mother (Null)</td>
<td>11</td>
<td>2</td>
<td>67.2</td>
<td>11.26</td>
<td>0.001</td>
<td>1.00</td>
</tr>
</tbody>
</table>

356 cm, respectively (Brough 2013). Minimum mother age at birth ranged between 8 and 24 yr.

Model fitting

Using standardised data (Gelman 2008), S.1 and S.3 global models easily converged, as did every model produced using the dredge function. Plots of model residuals vs. fitted values showed no indication of heteroscedasticity for either model set (Tables 1 & 2). Homogeneity of variance was also evident in plots of individual predictor parameters. The continuous predictor values of size and minimum age were both shown to be linear in relation to model predicted values, which suggests the chosen link function was appropriate (Bolker et al. 2009a). Assessing normality of model residuals and random conditional modes using qqnorm confirmed that our residuals and random effects followed a normal distribution for both model sets. Residual normality was further confirmed by Shapiro-Wilk tests that provided p-values of 0.14 and 0.28 (α = 0.05) for the S.1 and S.3 global models, respectively, indicating no significant departure from a normal distribution. The assumption that the random effects (see Fig. 4) approximated a normal distribution was supported for the S.3 analysis with the null (RE only) model producing a Shapiro-Wilk p-value of 0.07 (α = 0.05); however, the random conditional modes for the S.1 null model produced a Shapiro-Wilk p-value of 0.04, suggesting significant (α = 0.05) departure from a normal distribution. It has been suggested that the influence of non-normal modes is unlikely to compromise a GLMM fit except for extreme deviations from a normal distribution (Bolker 2015). The graphical representation of the S.1 conditional modes did not identify any extreme deviations from normal or outliers. Moreover, conditional modes from the S.1 global model were not significantly different from a normal distribution (p-value = 0.21). Thus, we assume that the slight deviation from normality for the S.1 conditional modes will not compromise our model fitting and subsequent parameter estimates (see Supplements 2 & 3 at www.int-res.com/articles/suppl/n029p255_supp.pdf for model diagnostic plots). Comparing residual deviance with residual degrees of freedom for both global models suggested that the data for each model set (Tables 1 & 2) were not overdispersed.

Survival to 1 yr

The most parsimonious S.1 model (indicated by the lowest AICc value and highest weight) combined month and mother size as predictors (Table 1). Mother size featured in all but one of the top models. The second best model, in which size was the only parameter, contributed 0.20 in model weight. Month of birth in particular featured strongly, with the models containing this parameter adding to a weight of 0.64 (Table 1). All other input parameters featured in the top S.1 models.
Although a ‘best’ model was selected, it did not have a model weight >0.9, and the second, third and fourth models contributed a reasonable amount to the cumulative weight index (Table 1). This indicates that no single model adequately explains variation in survival to 1 yr and confirms the utility of using model averaging to produce parameter estimations. Calf survival to 1 yr was significantly affected by month of birth and mother size (as indicated by 95% confidence intervals that did not overlap 0; Fig. 2). A positive coefficient of 1.75 indicates that calves born in January had a higher chance of survival. Moreover, it seems that calves born after January were also more likely to survive to 1 yr, although this effect is smaller (coefficient 0.71; Fig. 2). A model-averaged parameter estimate of the various factorial ‘Month’ levels were relative as the level ‘after Jan’ was selected as the intercept for model fitting. Tailrace and minimum mother age had positive mean values but their confidence intervals overlapped zero. Mother size had the highest model-averaged coefficient at 2.79. The only level of the month parameter with a significant effect on calf survival to 1 yr was January, with a model-averaged coefficient of 1.14 (Fig. 3).

Random effects

The influence of mother ID was slightly different for each model set (Table 3). In the S.1 analysis, the null model indicated strong variation in the conditional modes among individual females (Fig. 4). The overall model variance associated with the S.1 ran-
dom parameter was 0.84 for the null model. This variance was reduced for the S.1 global model (0.81) and top model (0.70; Table 3). This suggests that some among-individual variance in reproductive success is not accounted for by the fixed factors in the S.1 analysis. However, the S.1 null model occupied a low 12th rank in the overall model set and was not included in the top model set when ranked by AICc (Table 1).

The conditional modes from the S.3 null model suggest considerable variation among mothers in terms of calf survival to 3 yr (Fig. 2). This is further confirmed by overall model variation attributable to the S.3 null model random parameter of 0.591 (Table 3). The variance associated with the random parameter for the S.3 global and top models was 0. This suggests that for the S.3 analysis the total among group variance was accounted for by variation in the fixed factors. The S.3 null model also occupied a low position (11th) in the model selection table, contributing almost zero weight to the overall model set, and was not included in the top model set (Table 2). The low position of the S.3 null model and the 0 random variance term for the S.3 global and top models suggests that variation in reproductive success among females is reduced when the fixed factors are taken into consideration (Tables 2 & 3). That being said, it is still important to assess where differences in reproductive success exist; thus, the null model has been used to illustrate these differences (Fig. 4).

**Table 3. Variance associated with the random effects of mother ID for the null, global and top models for the survival to 1 (S.1) and 3 (S.3) yr analysis**

<table>
<thead>
<tr>
<th>Model</th>
<th>Random variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Null S.1</td>
<td>0.84</td>
</tr>
<tr>
<td>Global model S.1</td>
<td>0.82</td>
</tr>
<tr>
<td>Top model S.1</td>
<td>0.7</td>
</tr>
<tr>
<td>Null S.3</td>
<td>0.59</td>
</tr>
<tr>
<td>Global model S.3</td>
<td>0</td>
</tr>
<tr>
<td>Top model S.3</td>
<td>0</td>
</tr>
</tbody>
</table>

**Fig. 4. Conditional modes for the random effect of mother ID for the survival to 1 and 3 yr null models. Error bars are ± unconditional standard error**

**Top model set vs. simple models**

Using a simplifying rule to exclude complex models from the sets used for model averaging (Richards 2005, 2008) reduced the number of parameters for which estimates were available (Table 4). The simplifying rule made no difference to the parameters deemed to be statistically significant. Moreover, parameter estimates generated using the entire top model set were very similar to those from the simplified model set (Table 4). This indicated that model
averaging of input parameters across our top models sets was appropriate given the level of dispersion in our binomial response variable. That similar parameter estimates were produced using the top model and the stricter ‘Smodel’ set provides further validation of our results, as we arrived at the same conclusion using the 2 different model averaging methods.

**DISCUSSION**

**Modelling**

Using a mixed modelling approach combined with model averaging proved useful for unravelling the relative importance of the various predictor variables. However, there is often some bias associated with calf survival data due to unobserved birth events. Over the course of the study it is likely that some calving events were not documented (i.e. when a calf died shortly after birth). This problem is well recognised in other studies of calf survival in marine mammals (Mann et al. 2000, Kogi et al. 2004, Currey et al. 2009b). Failure to record a calving event in which the calf had died soon after birth would cause calf survival rates to be biased high. In contrast, overall calf production per year would be biased low (Henderson et al. 2014). In the present study, we assume that these biases will not influence model outcomes and subsequent parameter estimates. This assumption is reasonable as it is likely only a very small number of births attributable to the females in this study were missed. Tour operators in daily direct contact with the dolphins noted only 3 ‘potential’ births (followed by early mortality) that the monitoring team were unaware of. Seventy-four births were recorded between 1995 and 2012 (not all of these were included in the analysis as above). We consider 3 potentially missed births to be a sufficiently small proportion such that it would not affect our model outcomes, particularly as these birth events are by no means certain and may not have been attributable to the females in this study. In any case, these births were not included in the present study because it was not possible to establish the mothers’ identities.

Our exclusion of any mothers that could have been younger than 8 yr old minimised the likelihood of mother size changing significantly over the study period, as growth should have reached an assymp-
stronger in first year calf survival. The top model for the S.1 analysis still retained some variance associated with the random factor, demonstrating that not all variability in first year calf survival can be accounted for by the fixed effects. In contrast, RE variance for the S.3 analysis was effectively zero, suggesting that survival to 3 yr was completely accounted for by the fixed effects parameters. It has been discussed that zero RE variance may represent model over-fitting (Bolker 2015), which may bias random slope estimates for strongly unbalanced designs (Schielzeth & Forstmeier 2009). However, because of the binary response in our modelling framework, the zero RE variance estimate is likely to be consistent with a pure Bernoulli process with conditioning on the covariates.

Variability in reproductive success among individual bottlenose dolphin females has also been demonstrated in Shark Bay, Australia, where it has been attributed to inbreeding (Frère et al. 2010) and particular patterns of habitat use (Mann et al. 2000). Fruet et al. (2015) showed strong variation in female reproductive success in a bottlenose population in Brazil, where, as in Doubtful Sound, some individuals have never had a calf survive to 3 yr. High variance in reproductive success seems likely to be a feature of many dolphin populations.

**Mother size**

Calf survival to 1 and 3 yr was best explained by the size of the mother, with larger mothers being more successful (Figs. 2 & 3). Mother size is often correlated with her overall condition and reproductive fitness (Fairbanks & McGuire 1995, Derocher & Stirling 1998). In this study, as mother size was measured as length, and only in 2012; it does not accurately represent breeding condition, which is likely to change over time (Miller et al. 2011, 2012). Yet, the fact that variation in absolute body size among female dolphins in Doubtful Sound influences calf survival is an interesting and important result. Three main hypotheses have been advanced to explain how large female size may enhance the survival of offspring in mammals:

(1) Larger females may retain a higher position in the social hierarchy, presenting better opportunities to breed and potentially increasing exposure to top ranked males of superior fitness (Reiter et al. 1981, Wauters & Dhondt 1989).

(2) Larger females may give birth to larger offspring that are more capable of surviving the first years in a challenging environment (Atkinson & Ramsay 1995, Derocher & Stirling 1998).

(3) The energetic requirements of lactation are a significant tax upon the resources of mammalian mothers (Stewart & Lavigne 1984, Miller et al. 2012). If larger mothers have more resources, they may provide better quality and/or quantity of milk to their young, maximising chances of growth and survival (Iverson et al. 1993, Sakai & Harada 2001). However the opposite could also be true. Smaller mothers, not being required to support a large body size, may have greater resources available for the rearing of young (Pomeroy et al. 1999).

Bottlenose dolphins have a promiscuous mating system, with frequent mating (Scott et al. 1990, Whitehead & Mann 2000). In this context, the first hypothesis seems unlikely to be true. Hypotheses 2 and 3, however, seem to have more merit. In this study we were unable to assess whether larger females give birth to larger calves because photogrammetric sampling was carried out during 2012 and so was limited to the calves of that year only. For species inhabiting temperate and polar regions where food supply is variable and climatic conditions extreme, the benefits of maintaining a large body size for maximising reproductive success are likely to be important (Sand 1996, Derocher & Stirling 1998, Miller et al. 2012). Especially in marginal habitats, it may be that smaller females have fewer resources available to contribute to offspring (Sand 1996). It has also been suggested that smaller females may retain resources for themselves, sacrificing maternal investment to favour their own survival and potential for future reproduction (Fairbanks & McGuire 1995, Festa-Bianchet & Jorgenson 1998). In bighorn sheep this trend is particularly evident during periods of resource scarcity (Festa-Bianchet & Jorgenson 1998). As mother size had a significant influence on calf survival outcomes to both 1 and 3 years of age (Figs. 2 & 3), there is no evidence that large mother size benefits particular calf age classes in different ways.

**Birth timing**

Survival to 1 and 3 yr also seems to be strongly affected by the calf being born at particular times during the calving season (Figs. 2 & 3). Survival of calves to 1 yr was significantly predicted by being born during and after January, whilst January was the only significant month of birth level predicting survival to 3 yr. This suggests that calves born during and after January have a higher chance of survival.
relative to calves born during other periods. This trend was first identified in Doubtful Sound by Henderson et al. (2014), who showed that the 6 most successful mothers gave birth to a higher proportion (53%) of their surviving calves in January. The present study has confirmed this trend via statistical modelling of month of birth along with other potential predictor variables (Fig. 2).

Seasonal reproduction is the norm in a wide range of taxa including ungulates (Festa-Bianchet 1988, Gaillard et al. 1997), birds (Phillips et al. 1996, Vergara et al. 2007), reptiles (Moore et al. 1984) as well as marine mammals (Chittleborough 1958, Majluf 1992). Generally, bottlenose dolphin populations show narrow calving seasons at high latitudes (Mann et al. 2000, Haase & Schneider 2001), and broader calving seasons at low latitudes (Scott et al. 1990, Urian et al. 1996). Seasonal reproduction may have evolved to reduce climatic stresses upon offspring (Lims 1990, Ransome & McOwat 1994) and/or to take advantage of seasonally abundant resources for raising offspring (Daan et al. 1988, Goldizen et al. 1988). There are substantial seasonal fluctuations in water temperature and salinity within the Doubtful Sound complex (Gibbs et al. 2000). Moreover, cool water temperatures can persist well into the calving season as rainstorm events and cold fronts can cause dramatic decreases in water temperature (Gibbs et al. 2000). Cool water temperatures are likely to cause thermal stress to smaller, thinly insulated calves (Yeates & House 2008), a fact that may explain why calves that die in their first year typically die in the first month (Henderson et al. 2014).

It is possible that food resources within the fiord are seasonally limited, influencing the optimal timing for giving birth. Primary productivity in the New Zealand fiords is low due to low sunlight, limited space for macroalgae and high fluctuations in temperature and salinity (Peake et al. 2001, Tallis et al. 2004). Previous research has identified that dolphins in Doubtful Sound are more reliant on demersal, reef-dwelling prey species than on a seasonal influx of pelagic species (Lusseau & Wing 2006). Studies of other populations of bottlenose dolphins have linked seasonal reproduction to the availability of seasonal resources (Urian et al. 1996, Fruet et al. 2015). It may be that although pelagic species do not constitute a significant proportion of dolphin diet in Doubtful Sound, they are seasonally important for pregnant/lactating females during the calving season. Large schools of mackerel (Scomber australasicus and Trachurus declivis) can occur in the fiord during late spring and early summer (T. E. Brough pers obs.), thus, it is possible that the optimal January calving window is a trade-off between seasonally abundant resources during early summer (November/December) and the warmest water temperatures (February). Mean monthly temperature at birth was originally included in the analysis as a covariate but was subsequently excluded due to a correlation with month of birth.

Fruet et al. (2015) assessed female reproductive traits and success in a population of bottlenose dolphins inhabiting the sub-tropical southwest Atlantic. Substantial heterogeneity in female reproductive success was evident as was a narrow temporal ‘pulse’ for calf births during the calving season. It was suggested that the pulse in calf births may be related to warmer water temperatures and increased dolphin prey availability, which may favour calf survival during the pulse period. No evidence was provided to support this hypothesis, however.

There is some evidence for inter-annual variability in calf survival in this population (Currey et al. 2009b, Henderson 2013). Year-to-year variation in calf survival was not assessed in this study due to issues with the parameterisation of ‘year’ as a factor and subsequent problems with quasi- or complete separation (Abrahantes Cortiñas & Aerts 2012). Inter-annual variability in calf survival will be the focus of further study on this population (T. E. Brough et al. unpubl.).

**Tailrace**

The increase in freshwater inflow to the fiord system from the Manapouri hydro-electric scheme has caused the fiord’s low salinity layer to become deeper and colder at certain times of the year (Gibbs et al. 2000, Peake et al. 2001) and has dramatically altered some biological communities with the fiord (Tallis et al. 2004, Rutger & Wing 2006). These effects were most evident when the tailrace opened in 1971. The inclusion of a tailrace parameter, denoting the presence or absence of the second tailrace tunnel and its potential effects, was not shown to have a significant influence on calf survival in this study (Figs. 2 & 3). Due to the complicated relationship between the dolphins, their habitat and the tailrace, it may be that a binomial tailrace parameter was too simplistic to accurately describe variation in calf survival.

**Mother age**

Minimum mother age was not shown to be a significant predictor of calf survival to 1 or 3 yr (Figs. 2
Lack of accuracy in our parameterisation of age (i.e. underestimated age for females that were observed as adults at 1990) may have impacted our ability to establish age-related effects. Increasing age or experience may allow a mother to target her care towards her offspring at times when it is most needed (i.e. cooler temperatures; Green 1993, Cameron et al. 2000). This is the ‘targeted care hypothesis’, which proposes that while total reproductive effort may not increase with age, a mother’s ability to recognise important timing for application of maternal care may (Cameron et al. 2000). For long-lived species living in an extreme environment, such as the dolphins of Doubtful Sound, such application of maternal care may increase substantially a calf’s chances of survival. Age has been correlated with reproductive success in other bottlenose populations (Frère et al. 2010, 2015). As population monitoring continues in Doubtful Sound, age-related effects upon reproduction may be better determined in the future.

CONCLUSIONS

This study reveals how females within marine mammal populations can contribute unequally to the reproductive rate of the population, based on biological factors. In Doubtful Sound, heterogeneity in female reproductive success is attributable to larger mothers being more successful and the ability of females to give birth at the optimum time. These results further confirm that variable female reproductive success, in terms of calf survival, is the primary cause of vulnerability in this population. Future management of the population should focus on protecting vulnerable demographic groups. In particular, management could prevent or minimise future impacts that could influence the ‘sweet-spot’ in birth timing (i.e. alteration of environmental conditions) or minimise disturbance to calves during this time from tourism operations. These findings may also be important in an evolutionary context. It is well established that the bottlenose dolphins found at high latitudes are larger than those in warmer waters (Chong & Schneider 2001). Large mothers having higher reproductive success may well be the mechanism driving this phenomenon. For threatened populations or species that are limited by neonatal or juvenile survival, understanding the mechanisms that influence reproductive heterogeneity may prove crucial in developing management strategies to target anthropogenic contributions to vulnerability.

Acknowledgements. Firstly, we thank the numerous people that have contributed to the Doubtful Sound bottlenose dolphin photo identification dataset over the years: Jamie Williams, Karsten Schneider, Patti Haase, David Lusseau, Lucy Rowe, and Rohan Currey. This project was made possible via funding from the NZ Whale and Dolphin Trust, NZ Department of Conservation and the University of Otago. The team at the Fiordland district office of the Department of Conservation have been invaluable to this project in the last decade. We particularly thank Dave Johnson, Richard Kinsey, Pete Young and Dallas McHardy. Thanks to Real Journeys Ltd. and Deep Cove education trust for transport and accommodation in the field.

LITERATURE CITED


Brough TE (2013) Using photography to study the conservation biology of bottlenose dolphins in southern New Zealand. MSc thesis, University of Otago, Dunedin


of brown trout affected by copper and zinc. Environ Toxicol Chem 34:816–820


Editorial responsibility: Robert Harcourt, Sydney, NSW, Australia

Submitted: July 17, 2015; Accepted: November 30, 2015
Proofs received from author(s): January 21, 2016