



**Report on an estimation model for the
pupping rate of New Zealand sea lions**

**NIWA Client Report: WLG2008-03
January 2008**

NIWA Project: DOC08302

Report on an estimation model for the pupping rate of New Zealand sea lions

Author
David Gilbert

Prepared for

Department of Conservation Project

POP2006-01 Objective 3. Analysis from sea lion database to estimate pupping rate and associated parameters

NIWA Client Report: WLG2008 03
January 2008

NIWA Project: DOC08302

National Institute of Water & Atmospheric Research Ltd
301 Evans Bay Parade, Greta Point, Wellington
Private Bag 14901, Kilbirnie, Wellington, New Zealand
Phone +64-4-386 0300, Fax +64-4-386 0574
www.niwa.co.nz

Contents

1. Introduction	1
2. Model	1
3. Data	3
4. Estimable parameters	3
5. Pupping rate and natural mortality functions	4
6. Likelihoods	5
7. Discussion	8
References	9

Approved for release by:

John McKoy
General Manager, Fisheries Research

Executive summary

Here we report the second stage in this project. Mathematical formulae are developed to allow the estimation of pupping rate for New Zealand sea lions from tag-resighting data. The method involves defining functions and their associated parameters that measure the likelihood of the behavioural observations. These functions are multiplied across all observations to produce total likelihood. A pupping rate function is defined and estimated, based on several parameters. A mortality function must also be defined to allow the pupping rate to be estimated. The method involves treating the observations as arising from mixtures of statistical distributions. This is necessary because not all animals can be identified as definite breeders or definite non-breeders each season. Our initial report (Gilbert 2007) indicated that cows could be placed into two categories, high- and low-fecundity animals and parameters will be estimated to account for this. In the third stage of this project we will minimise the total likelihood with respect to the parameters. Our primary result will be a smooth, domed function giving the pupping rate at age for female New Zealand sea lions that breed at Enderby Island. It will be a weighted average of functions for high- and low-fecundity animals where the weightings will be the estimated proportions in each group.

1. Introduction

Our initial report (Gilbert 2007) discussed the viability of estimating pupping rate for New Zealand sea lions from the available tag-resighting data. It concluded that good estimates were possible. Various features of the data were examined. Preliminary pupping rate estimates were presented and will not be repeated here. The present report documents the next step in the project, which is the development of a mathematical model to estimate pupping rate. Our goal is to describe pupping rate as a function of age and to estimate a set of parameters that determines the function. To allow this, a number of other parameters must also be estimated that describe observation and mortality processes. A function of age is developed to describe annual mortality. Then the likelihood formulae are developed that express the probabilities of making the observations that were actually made, conditional on the parameters. These are defined. The hypotheses underlying these likelihoods are discussed. Modelling of this sort is a process in which the steps: hypothesis testing, diagnosis and hypothesis reformulation, are repeated until satisfactory diagnostics are obtained. Some of the possible directions of hypothesis reformulation are mentioned. Our expectation is that a satisfactory model will emerge from which a pupping rate function will be estimated and that this will be broadly similar to the preliminary estimate but will deviate from it in ways discussed by Gilbert (2007).

2. Model

To estimate pupping rate, a statistical model is required to explain the variability in the numbers of behavioural observations. The model is governed by a set of parameters, which are estimated by maximum likelihood. A satisfactory model will fit both the observation means and the patterns of variability around those means.

It is desirable to reduce the number of estimable parameters by expressing relationships in functional forms dependent on a few parameters. Over-parameterisation leads to estimates that lack robustness and tend to show random variations in the data rather than underlying relationships. The preliminary estimate of pupping rate in Gilbert (2007) has implausible spikes at ages 9, 12, 16 and 19 years. Here we express both mortality and pupping rate as smooth functions of age. The functions are parameterised so that the parameters are meaningful. For example, β_1 is the maximum pupping rate and it occurs for cows aged β_2 years.

Because adequate behavioural observations commenced in 2000, the only data used prior to that season are the tagging season and the recorded deaths. Positively identified deaths of tagged cows recorded outside the breeding season are included. Apart from these the model is fitted only to the 2000–2007 breeding season data.

The breeding season behavioural observations are categorised into “Breeder”, “Probable breeder”, “Possible breeder” and “Unknown”. This gives a vector, the first element of which is binary (0 or 1) and the other elements are counts. This data vector may be all zeros for a given season if the cow is dead or if it is alive but not sighted during that season. If a cow is sighted in a subsequent season the first alternative can be ruled out of the analysis.

The resighting effort is similar amongst seasons because the field trips are in most seasons about 85 days and are carried out according to a fixed routine. However the estimable parameter ϱ is used to account for any variability due to differences in effective resighting effort. This could be caused by differences in length of trip,

weather, size of team and extent of other activities undertaken. Whether this parameter significantly explains variance or correlates with trip length will be tested.

The Poisson distribution is a simple statistical explanation for behavioural counts. All non-overlapping time periods (or amounts of resighting effort) of equal size have equal probabilities of containing an observation and these probabilities are independent of each other. The preliminary analyses suggested that the counts are over-dispersed, i.e. contain more very high and very low counts than would occur under the Poisson distribution. We will attempt to develop a good statistical explanation of the data based on the preliminary analyses. The behavioural observations come from a mixture of breeders and non-breeders. This can be characterised by a mixture of statistical distributions which will have a greater dispersion than if a single distribution were used. Except for observations where breeding can be inferred with certainty, the pupping rate (at age) determines the proportions of breeders and non-breeders in the mixture. The model must estimate both the pupping rate to determine the proportions in the mixture and the distributions that apply to each part of the mixture. Observations of cows nursing a pup contain many zeros and a non-zero distribution that is slightly over-dispersed (Figure 1). The zeros are largely for the non-breeders and the non-zeros the breeders. The frequency of non-zero observations of a cow with a pup show greater over-dispersion (Figure 2). Some non-zero counts for non-breeders are likely and so the distribution may perhaps be explained by a mixture of Poisson distributions of breeders and non-breeders. Observations of cows exhibiting no behaviour relevant to breeding are considerably over-dispersed (Figure 3). It appears that a mixture of Poisson distributions of breeders and non-breeders with different mean season frequencies would be insufficient to adequately explain this dispersion. We therefore propose to use the negative binomial distribution, which allows the variance of the counts to be a multiple of the mean (not equal to it as in the Poisson distribution). The negative binomial distribution could arise from heterogeneity in the behavioural observation means amongst both breeders and non-breeders. We will assume that the behavioural counts in each part of the mixture are distributed according to the negative binomial distribution.

So far we have assumed that whether a cow breeds in a season depends only on its age, not on its previous breeding history. We might expect that the physiological burden of bearing a pup would reduce the probability of a cow bearing a pup in the following season. The data does not support this hypothesis (Figure 4). Cows are more likely to bear a pup in a given season if they bore one in the previous season. It appears instead that some cows exhibit high fecundity by breeding frequently, whereas others breed only occasionally. The sequences of observations for individuals support this hypothesis. We will therefore assume that some proportion of the population (to be estimated) is highly fecund. The low-fecundity females will have a lower pupping rate function than those of high fecundity and both will be estimated.

It also appears that some cows are inherently more readily resighted than others, with repeated large resighting counts. Both individual behaviour, the early loss of the tag from one flipper and tag readability could contribute to this. We will therefore assume that some proportion of the population (to be estimated) is highly resightable. Again the low-resightability animals will have lower mean observation counts than those of high resightability and these will be estimated.

These assumptions result in mixtures of negative binomial distributions that we hope will be able to fit these highly dispersed data.

3. Data

Let I , a , i , and y denote tag number identity, age, behaviour category and season. The data will be arranged in the following format:

1. The number of tagged cows, N .
2. A set of N blocks of data, one for each tagged cow. The first row in each block contains: tag number I , tag season y_T , the last resight season, y_L and the observed season of death, y_D . For cows where there have been no observations after 1999, y_L is null. y_D will also frequently be null because few animals are observed to have died either during the breeding season or in a fishery. y_L is required because in the seasons following y_L the possibilities of both death and survival without sighting must be included in the likelihood formulae. If $y_L \geq 2000$, this is followed by $y_L - 2000 + 1$ rows of observations, one for each season from 2000 to y_L . Each row has four columns, one for each behaviour grouping. The first behaviour grouping refers to the observations, BIRTH, DEADPUP, PREGNANT, and STILLBIRTH. Mostly these behaviours are made once per season and indicate definite breeding. If at least one such behaviour is observed, the column will contain a one and otherwise a zero. The second grouping refers to only the behaviour NURSE, which indicates that breeding is highly probable and the column contains the number of such observations. The third grouping refers to the behaviours CALL and WITHPUP, which indicate that breeding is probable and the column contains the combined number of such observations. The fourth grouping refers to the behaviours, DEAD, X (no relevant behaviour), XSUCKLING (suckling from a cow) and YNURSE (nursing a yearling), which do not indicate breeding but do not necessarily exclude it. The column contains the combined number of such observations (in reality YNURSE does exclude breeding but the frequency of this behaviour is so low that this fact has been ignored).

For a given tag number in season y , the cow is of age $y - y_T + 1$ and the vector of behavioural observations is denoted by $\underline{x}_y = (x_{y1}, x_{y2}, x_{y3}, x_{y4})$.

4. Estimable parameters

The estimable parameters which determine the likelihoods of the behavioural observations are given here. Bounds indicate anticipated values.

We define pupping rate as a smooth, domed function of age that gives the probability that a high-fecundity cow will bear a pup in a given season. The function is determined by a vector of parameters, $\underline{\beta}$, where β_1 is the maximum value of the function, $0.4 < \beta_1 < 0.95$, β_2 is the age at which the maximum is achieved, $8 < \beta_2 < 14$, β_3 and β_4 are the ages at which the left and right-hand limbs of the function fall to half the maximum, $4 < \beta_3 < 8$ and $\beta_4 > 14$,

we define mortality as a smooth, U-shaped function of age that gives the probability that a cow will die in the year after achieving a given age, where μ_1 is the minimum function value (mortality rate), $0 < \mu_1 < 0.3$, μ_2 is the age at which the minimum is achieved, $0 < \mu_2 < 8$, μ_3 is the function value at age 0, $0 < \mu_3 < 0.8$, and μ_4 is the function value at age 20, $0 < \mu_4 < 0.8$,

let ϕ be the probability that a given cow is of low fecundity, $0.2 < \phi < 0.8$,

let ρ be the ratio of pupping rates of low-fecundity cows to that of high fecundity cows, $0.1 < \rho < 0.5$,

- let ψ be the probability that a given cow is of low resightability, $0.2 < \psi < 0.8$,
- let τ be the ratio of expected numbers of sightings of low-resightability, breeding cows to that of high resightability breeding cows, $0.2 < \tau < 0.8$,
- let $\underline{\pi}$ be the vector of length 4 where π_1 is the probability of making a “Breeder” observation and π_2, π_3 and π_4 are the expected numbers of sightings of “Probable breeder”, “Possible breeder” and “Unknown” behaviour groupings for a high resightability cow in a season that it breeds,
- let $\underline{\pi}'$ be the equivalent to $\underline{\pi}$ for a high resightability cow in a season that it does not breed. π'_1 is fixed at 0 because a “Breeder” observation is impossible for a cow that it does not breed. It is not estimated but is retained for consistency with $\underline{\pi}$,
- let $\underline{\theta}$ be the vector of length 8 of the effective resighting effort for the seasons 2000 – 2007, where θ_{2000} is set at one and the other values are multipliers,
- let V be the proportionality constant that multiplies the numbers of expected behavioural observations to give the variance under the negative binomial model.

5. Pupping rate and natural mortality functions

The pupping rate function expresses the probability that a cow of a given age will bear a pup in a given season, and therefore includes the process of maturity. We therefore expect it to have a rising left hand limb, to reach a maximum between 8 and 14 years and then to decline gradually with age. This is the empirical shape found by Chilvers (2006) and Gilbert (2007). Gilbert (2007) and Figure 4 suggest that some cows exhibit high fecundity and breed in most seasons following maturity, whereas others only breed occasionally. We therefore divide the population into high- and low-fecundity groups. The function we define depends on parameter vector $\underline{\beta}$ and refers to the high-fecundity group,

$$b(a | \underline{\beta}) = \begin{cases} \beta_1 2^{-\left(\frac{a-\beta_2}{\beta_3-\beta_2}\right)^2} & a \leq \beta_2 \\ \beta_1 2^{-\left(\frac{a-\beta_2}{\beta_4-\beta_2}\right)^2} & a > \beta_2 \end{cases}$$

The function has a maximum value of β_1 at $a = \beta_2$, takes the value $0.5\beta_1$ at $a = \beta_3$ on the left and at $a = \beta_4$ on the right.

If the high- and low-fecundity animal hypothesis is supported by the data then an average pupping rate for the population can be obtained by taking a weighted average of the respective pupping rates. The weightings are the estimated proportions in each group,

$$(1 - \phi) b(a | \underline{\beta}) + \phi \rho b(a | \underline{\beta})$$

Estimation of this function is essentially the goal of the project. However other functions must be estimated to achieve this.

We require a mortality rate for each age class to determine how many cows that were not resighted in 2007 were still alive in each prior season that they were not resighted. We will define mortality as the probability that a cow dies between 1 December and 30 November. Mortality is expected to follow an asymmetric U-shape, typical of natural mortality in vertebrates, with high mortality for the youngest animals, declining steeply to a minimum and then increasing gradually. Mortality here includes

incidental bycatches in fisheries (largely squid trawling) but also tag loss and total loss of tag readability. These other processes could potentially be separately estimated but are outside the scope of this project and are not essential to it. Because these other rates are relatively small and the preliminary estimate of mortality was U-shaped (Gilbert 2007), we assume that all these processes can be accommodated by the following function. It depends on parameter vector $\tilde{\mu}$,

$$m(a | \tilde{\mu}) = \begin{cases} \mu_1 \left(\frac{\mu_3}{\mu_1} \right)^{\binom{a-\mu_2}{0-\mu_2}} & a \leq \mu_2 \\ \mu_1 \left(\frac{\mu_4}{\mu_1} \right)^{\binom{a-\mu_2}{20-\mu_2}} & a > \mu_2 \end{cases}$$

The function has a minimum value of μ_1 at age $a = \mu_2$, takes the value μ_3 at $a = 0$ and μ_4 at $a = 20$.

6. Likelihoods

The method of maximum likelihood obtains parameter estimates by finding the set of values that make the whole set of observations the most likely. The likelihood function must be expressed algebraically to achieve this. In simple terms, the likelihood function is the product of the probabilities of each of the separate observations. This can be built up from its constituent parts. We first use the Poisson distribution, which is the simplest way of modelling counts. For the ‘‘Breeder’’ category, with values one or zero, the likelihood is either π_1 or $(1 - \pi_1)$. Given that a cow is alive at the start of season y and its fecundity, resightability and breeding status is known, the likelihood of the observation vector, \underline{x}_y , depends on the parameter vector, $\underline{\pi}$, and the effective resighting effort, θ_y ,

$$f(\underline{x}_y | \underline{\pi}, \theta_y) = (\theta_y \pi_1)^{x_{y1}} (1 - \theta_y \pi_1)^{1-x_{y1}} \prod_{i=2}^4 \frac{(\theta_y \pi_i)^{x_{yi}}}{x_{yi}!} e^{-\theta_y \pi_i}$$

For a cow that does not breed in season y , $x_{y1} = 1$ is impossible and results in $f(\underline{x}_y | \underline{\pi}', \theta_y) = 0$. The preliminary analyses suggest that the behaviour counts are over-dispersed relative to the Poisson distribution. The negative binomial distribution can account for over-dispersion using a parameter, K_{yi} , which depends on the parameter V . Given that a cow is alive at the start of season y , the likelihood of the observation vector, \underline{x}_y , depends on the parameter vector, $\underline{\pi}$, and the effective resighting effort, θ_y ,

$$f(\underline{x}_y | \underline{\pi}, \theta_y, V) = (\theta_y \pi_1)^{x_{y1}} (1 - \theta_y \pi_1)^{1-x_{y1}} \prod_{i=2}^4 \frac{\Gamma[K_{yi} + x_{yi}]}{\Gamma[x_{yi} + 1] \Gamma[K_{yi}]} \left(\frac{\theta_y \pi_i}{\theta_y \pi_i + K_{yi}} \right)^{x_{yi}} \left(1 + \frac{\theta_y \pi_i}{K_{yi}} \right)^{-K_{yi}}$$

As before, for a cow that does not breed in season y , $x_{y1} = 1$ results in $f(\underline{x}_y | \underline{\pi}', \theta_y, V) = 0$. The over-dispersion parameter of category i counts in season y , K_{yi} is a function of V ,

$$K_{yi} = \frac{\theta_y \pi_i}{V - 1}, \quad V > 1$$

The variance $\sigma_{y_i}^2$ is given by,

$$\sigma_{y_i}^2 = V\theta_y\pi_i$$

The over-dispersion of the behavioural observations is a constant multiple of the mean. The data will first be fitted using the Poisson distribution, where effectively $V = 1$, then with the negative binomial, and then with the negative binomial with different over-dispersion multipliers for different behaviour groupings, until satisfactory diagnostics are obtained.

We obtain the unconditional likelihood by combining the likelihoods of high- and low-fecundity cows, in proportions $(1 - \phi)$ and ϕ , with the low-fecundity pupping rates adjusted by the factor, ρ . We also combine the likelihoods of high- and low-resightability cows in proportions $(1 - \psi)$ and ψ , with the low-resightability expected counts adjusted by the factor, τ . Hence, given that a cow is alive at the start of season y , the likelihood of the observation vector, \underline{x}_y depends on the parameters, $\underline{\pi}$, $\underline{\pi}'$, θ_y , ϕ , ρ , ψ , τ , $\underline{\beta}$ and V ,

$$g(\underline{x}_y | \underline{\pi}, \underline{\pi}', \theta_y, \phi, \rho, \psi, \tau, \underline{\beta}, V) = \phi \left[\begin{aligned} & \psi \rho b(y - y_0 + 1 | \underline{\beta}) f(\underline{x}_y | \tau \underline{\pi}, \theta_y, V) + \\ & (1 - \psi) \rho b(y - y_0 + 1 | \underline{\beta}) f(\underline{x}_y | \underline{\pi}, \theta_y, V) + \\ & \left(1 - \rho b(y - y_0 + 1 | \underline{\beta}) \right) f(\underline{x}_y | \underline{\pi}', \theta_y, V) \end{aligned} \right] \\ + (1 - \phi) \left[\begin{aligned} & \psi b(y - y_0 + 1 | \underline{\beta}) f(\underline{x}_y | \tau \underline{\pi}, \theta_y, V) + \\ & (1 - \psi) b(y - y_0 + 1 | \underline{\beta}) f(\underline{x}_y | \underline{\pi}, \theta_y, V) + \\ & \left(1 - b(y - y_0 + 1 | \underline{\beta}) \right) f(\underline{x}_y | \underline{\pi}', \theta_y, V) \end{aligned} \right]$$

The effect of high and low resightability is assumed to be exhibited only in the seasons an animal breeds, but this assumption could be reversed if the data do not support it. In what follows the notation $g(\underline{x}_y | \dots)$ is used to avoid repeatedly specifying the parameters.

The total likelihood is obtained by multiplying together the likelihood from each season's observations. For a cow that was alive in 2000 but for which $y_T < 2000$, it is necessary to calculate the likelihood that it has survived to year 1999 using the mortality function,

$$L_{1999} = \prod_{y=y_T+1}^{1999} \left[1 - m(y - y_T - 1 | \underline{\mu}) \right]$$

Behavioural observations are only used from 2000 onwards. The likelihood of observations up to season y_L , if $y_L \geq 2000$, must include the probabilities of survival,

$$L_{y_L} = \prod_{y_1=y_T+1}^{1999} \left[1 - m(y_1 - y_T - 1 | \underline{\mu}) \right] \times \prod_{y_2=y_{2000}}^{y_L} \left[\left(1 - m(y_2 - y_T - 1 | \underline{\mu}) \right) g(\underline{x}_{y_2} | \dots) \right]$$

If $y_D \leq 2007$, i.e. y_D is known, then the likelihood of observations up to 2007 equals the likelihood up to season y_D ,

$$L_{2007} = \prod_{y_1=y_T+1}^{1999} \left[1 - m(y_1 - y_T - 1 | \underline{\mu}) \right] \times \prod_{y_2=y_{2000}}^{y_D} \left[\left(1 - m(y_2 - y_T - 1 | \underline{\mu}) \right) g(\underline{x}_{y_2} | \dots) \right] \times m(y_D - y_T | \underline{\mu})$$

If $y_D \leq 2007$, i.e. y_D is known, then the likelihood of observations up to 2007 equals the likelihood up to season y_D ,

$$L_{2007} = \prod_{y_1=y_T+1}^{1999} [1-m(y_1 - y_T - 1 | \mu)] \times \prod_{y_2=2000}^{y_D} [(1-m(y_2 - y_T - 1 | \mu)) g(x_{y_2} | \dots)] \times m(y_D - y_T | \mu)$$

If y_D is unknown and $y_L \geq 2000$, i.e. there are no behavioural observations 2000 – 2007, it is necessary to sum all the alternative seasons in which mortality may have occurred, i.e. y_T to 2007,

$$L_{2007} = m(0 | \mu) + \sum_{y_1=y_T+1}^{1999} \left\{ \prod_{y_2=y_T+1}^{y_1} [1-m(y_2 - y_T - 1 | \mu)] \times m(y_1 - y_T | \mu) \right\} + \sum_{y_3=2000}^{2007} \left\{ \prod_{y_4=2000}^{y_3} [(1-m(y_4 - y_T - 1 | \mu)) g(0_{y_4} | \dots)] \times m(y_3 - y_T | \mu) \right\}$$

where 0_y denotes a behavioural observation vector of zeros in season y . If y_D is unknown and $y_L \geq 2000$,

$$L_{2007} = \prod_{y_1=y_T+1}^{1999} [1-m(y_1 - y_T - 1 | \mu)] \times \prod_{y_2=2000}^{y_1} [(1-m(y_2 - y_T - 1 | \mu)) g(x_{y_2} | \dots)] \times \sum_{y_3=y_L+1}^{2007} \left\{ \prod_{y_4=y_L+1}^{y_3} [(1-m(y_4 - y_T - 1 | \mu)) g(0_{y_4} | \dots)] \times m(y_3 - y_T | \mu) \right\}$$

The last factor is omitted if $y_L = 2007$. Otherwise it includes the possibilities that the cow is dead and that it survives but is not observed in each season from $y_L + 1$ to 2007. In the above it is assumed that $y_T < 2000$. If $y_T \geq 2000$, y_D is unknown and $y_L \geq 2000$,

$$L_{2007} = g(x_{y_T} | \dots) \times \prod_{y_1=y_T+1}^{y_1} [(1-m(y_1 - y_T - 1 | \mu)) g(x_{y_1} | \dots)] \times \sum_{y_2=y_L+1}^{2007} \left\{ \prod_{y_3=y_L+1}^{y_2} [(1-m(y_3 - y_T - 1 | \mu)) g(0_{y_3} | \dots)] \times m(y_2 - y_T | \mu) \right\}$$

The other alternatives expressions for L_{2007} when $y_T \geq 2000$ are obtained similarly.

Estimation is achieved by summing the negative logarithms of the likelihoods of the observations over tag numbers, I ,

$$\Lambda = -\sum_I \ln(L_{2007}^I)$$

and minimising Λ with respect to the 28 estimable parameters. Modifications to the model will be made successively until satisfactory diagnostics are achieved.

We propose to carry out this minimisation using AD Model Builder© (Otter Research Ltd, Nanaimo, Canada), a powerful minimisation tool. There is some uncertainty as to whether the expression for $g(x_y | \pi, \pi', \theta_y, \phi, \rho, \psi, \tau, \beta, V)$, which involves a mixture of six distributions will allow parameter estimation. Good starting values for the parameters may be necessary. The preliminary report (Gilbert 2007) can provide some of these. If estimation fails, it may be necessary to drop the high- and low-resightability categories to reduce the mixture to four distributions.

7. Discussion

A feature of the method discussed above, in which likelihoods are defined as mixtures of statistical distributions, is that individual animals are not identified as being of high- or low-fecundity or as being of high- or low-resightability. These characteristics are expressed only as probabilities. Similarly, for animals not observed dead and not seen in 2007 we only estimate a probability that they are alive. Nor are the breeders and the non-breeders fully identified each season. The only identified breeders are those for which a “Breeder” observation has been made. These uncertainties are in fact the virtue of this method. The observations of some animals in some seasons would not allow their certain classification as breeders or non-breeders, and fecundity, resightability and death will generally be uncertain. All animals must nevertheless be accounted for and the method described does this by using probabilities. Because it is of interest, we will however calculate odds ratios as an indication of the probable categories the animals fall into. For each individual we will calculate its high-fecundity odds ratio i.e. the likelihood that it is of high fecundity over the likelihood that it is of low fecundity. Similarly we will calculate its high-resightability odds ratio. For each year we will also calculate the odds ratio that the cow was alive and bore a pup. Thus we can compare the results from the model with the criterion-based approach to identifying breeders that has been used previously by Chilvers (2006) and Gilbert (2007). In the criterion-based approach, observations that indicate probable breeding are used to identify breeders, but there is an unavoidable level of arbitrariness in classifying some animals. The model results may help identify strengths and weaknesses in the criterion-based method and may allow the definition of a better criterion for year-to-year usage.

In developing the model the primary tool we will use for determining whether a modification should be included is the AIC (Akaike Information Criterion). The following modifications will be considered. The 1998 cohort, which suffered high pup mortality in an epizootic, showed lower fecundity as adults (Gilbert 2007). A fecundity parameter for this cohort would probably significantly improve the model fit. Therefore cohort-specific and annual fecundity factors will be tested for inclusion. Likewise cohort-specific and annual mortality factors will be tested. Variation of the negative binomial variance multiplier, V , amongst observation types will also be tested. We will test whether there is evidence for different pupping rate functions and different mortality functions for high- and low-fecundity cows. If the high- and low-resightability hypothesis is supported we will test whether it applies to cows both in the seasons that they breed and in the non-breeding seasons. As with all model development, the results will depend on whether a model can be found that fits the data with satisfactory diagnostics.

The model we have described will provide parameter estimates for a smooth, domed function giving the pupping rate at age for female New Zealand sea lions that breed at Enderby Island. It will, most likely, be derived as a weighted average of the pupping rate functions for high- and low-fecundity animals. The weightings will be the estimated proportions in each group. We will test whether the pupping rate function for low-fecundity females is simply a scaled version of that for high-fecundity females or whether it has a different shape. Our hope is that as well as estimating pupping rate the model will improve our understanding of the New Zealand sea lion population.

References

- Gilbert, D.J. 2007: Report on the viability of pupping rate estimation for New Zealand sea lions. Report for Department of Conservation Project POP2006-01 Objective 3. NIWA Client Report: WLG2007-83, November 2007. 12 p.
- Chilvers, L. 2006. NZ sea lion research trip, Auckland Islands, November 20th 2005 to February 20th 2006. Department of Conservation Internal Report. May 2006. 19 p.

NURSE observations, ages 8 to 14 y Zero-censored Poisson fit

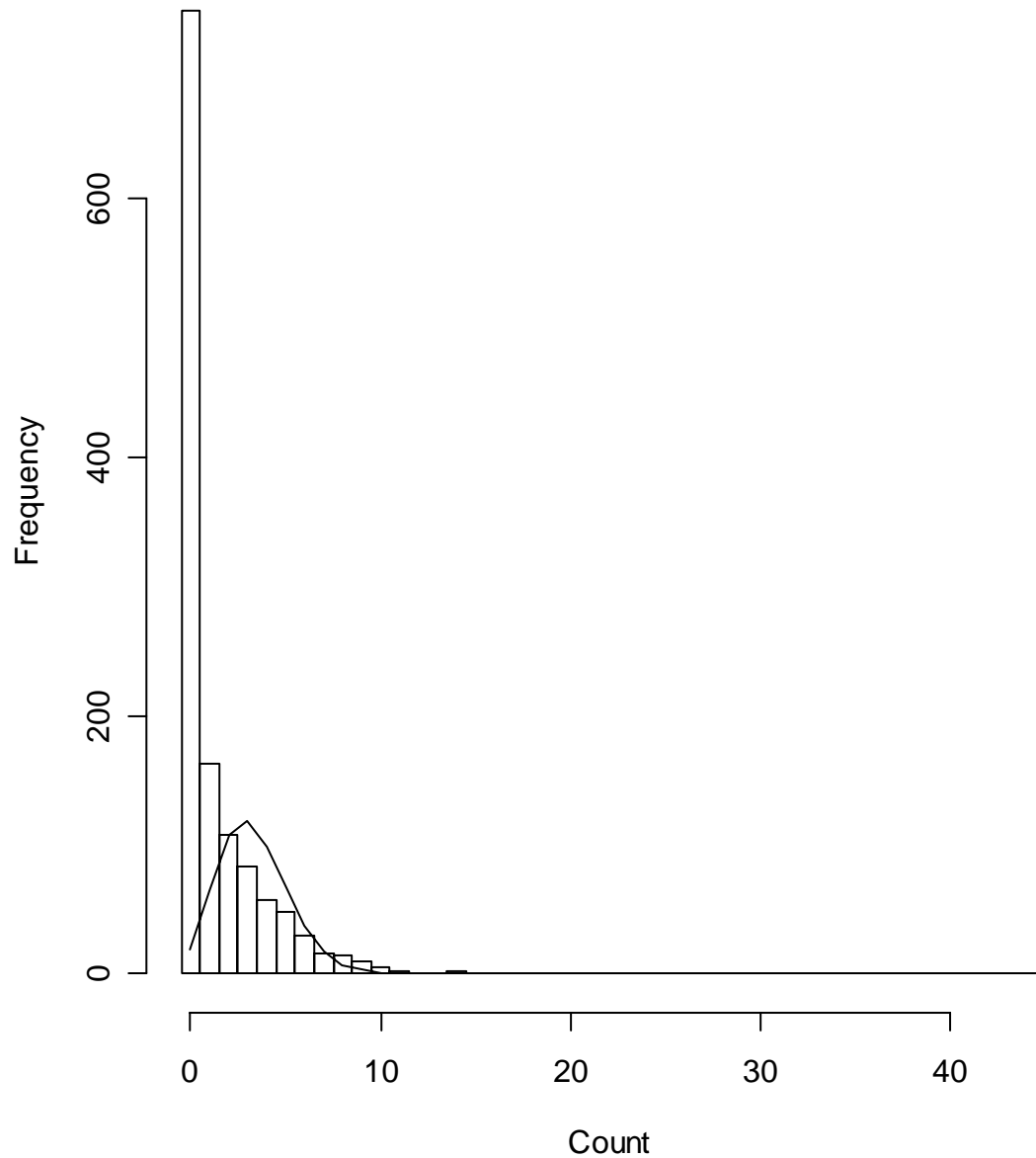


Figure 1. Frequency histogram of NURSE observations per season for tagged cows known to be alive based on Enderby Island observations after 1999. Only animals of ages 8 – 14 years are included. The line is fitted assuming a Poisson distribution after censoring the zero counts.

WITHPUP observations, ages 8 to 14 y Zero-censored Poisson fit

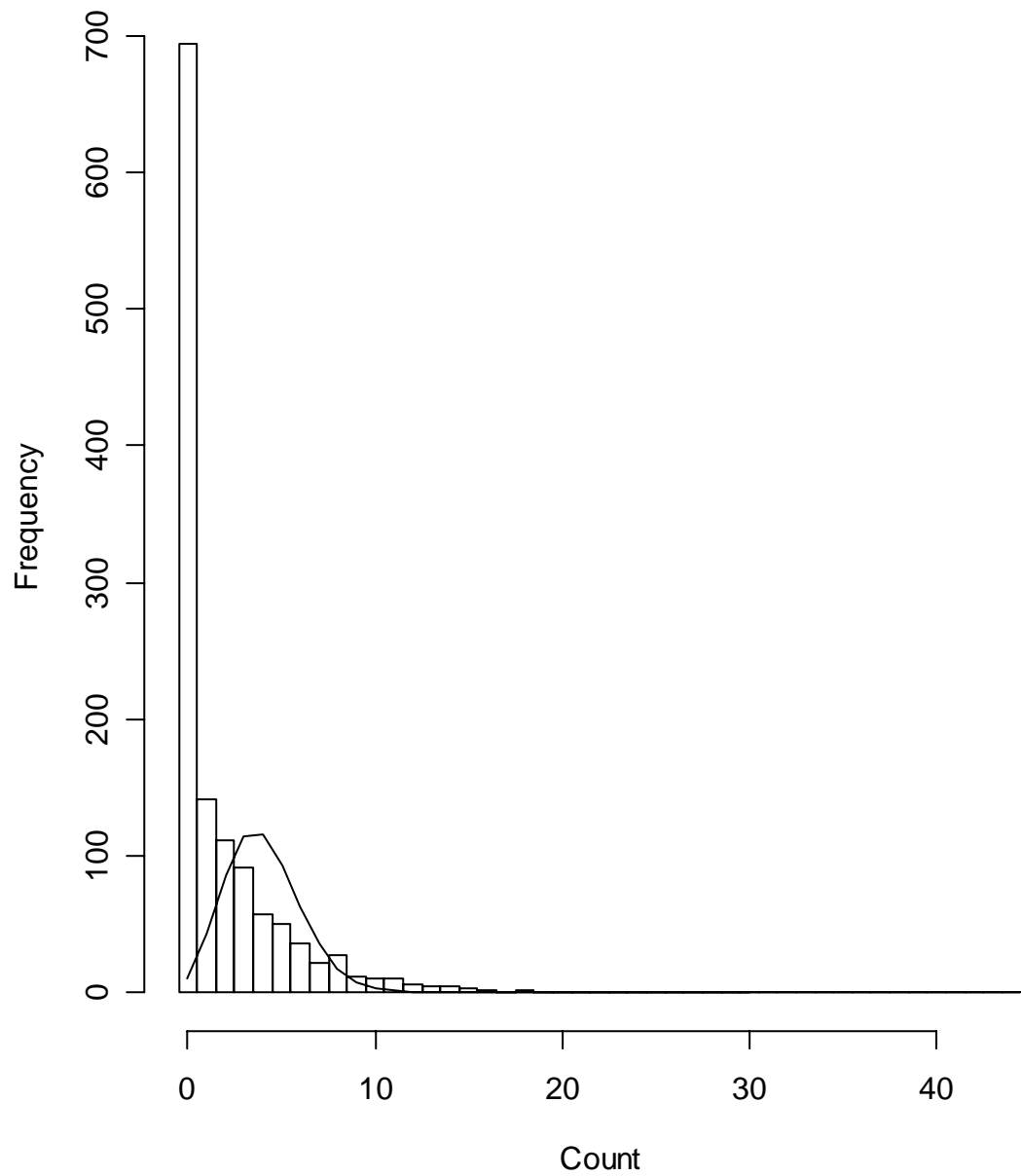


Figure 2. Frequency histogram of WITHPUP observations per season for tagged cows known to be alive based on Enderby Island observations after 1999. Only animals of ages 8 – 14 years are included. The line is fitted assuming a Poisson distribution after censoring the zero counts.

**X observations, ages 8 to 14 y
Poisson fit**

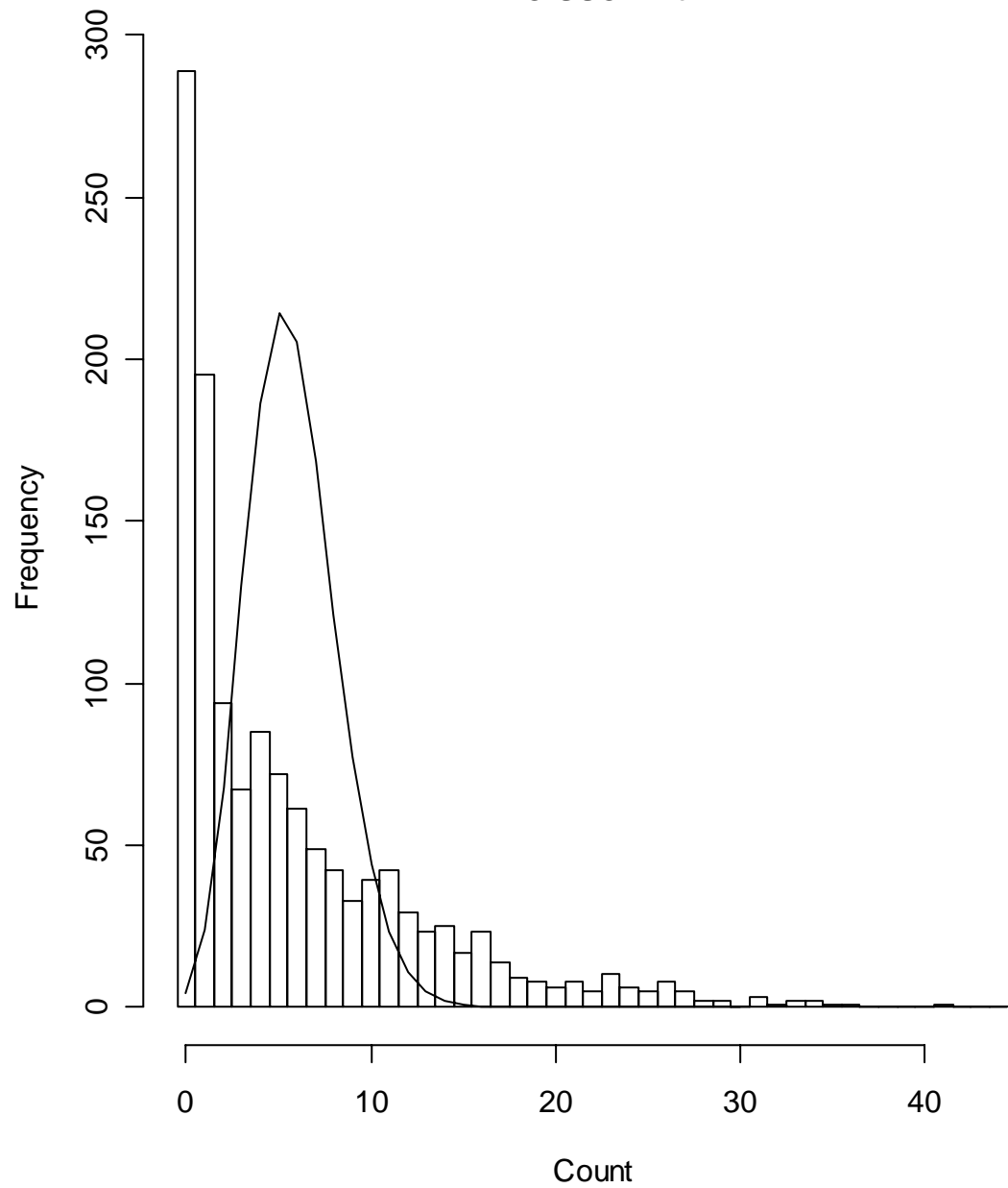


Figure 3. Frequency histogram of X observations (no behaviour relevant to breeding) per season for tagged cows known to be alive based on Enderby Island observations after 1999. Only animals of ages 8 – 14 years are included. The line is fitted assuming a Poisson distribution.

Proportion of females that breed

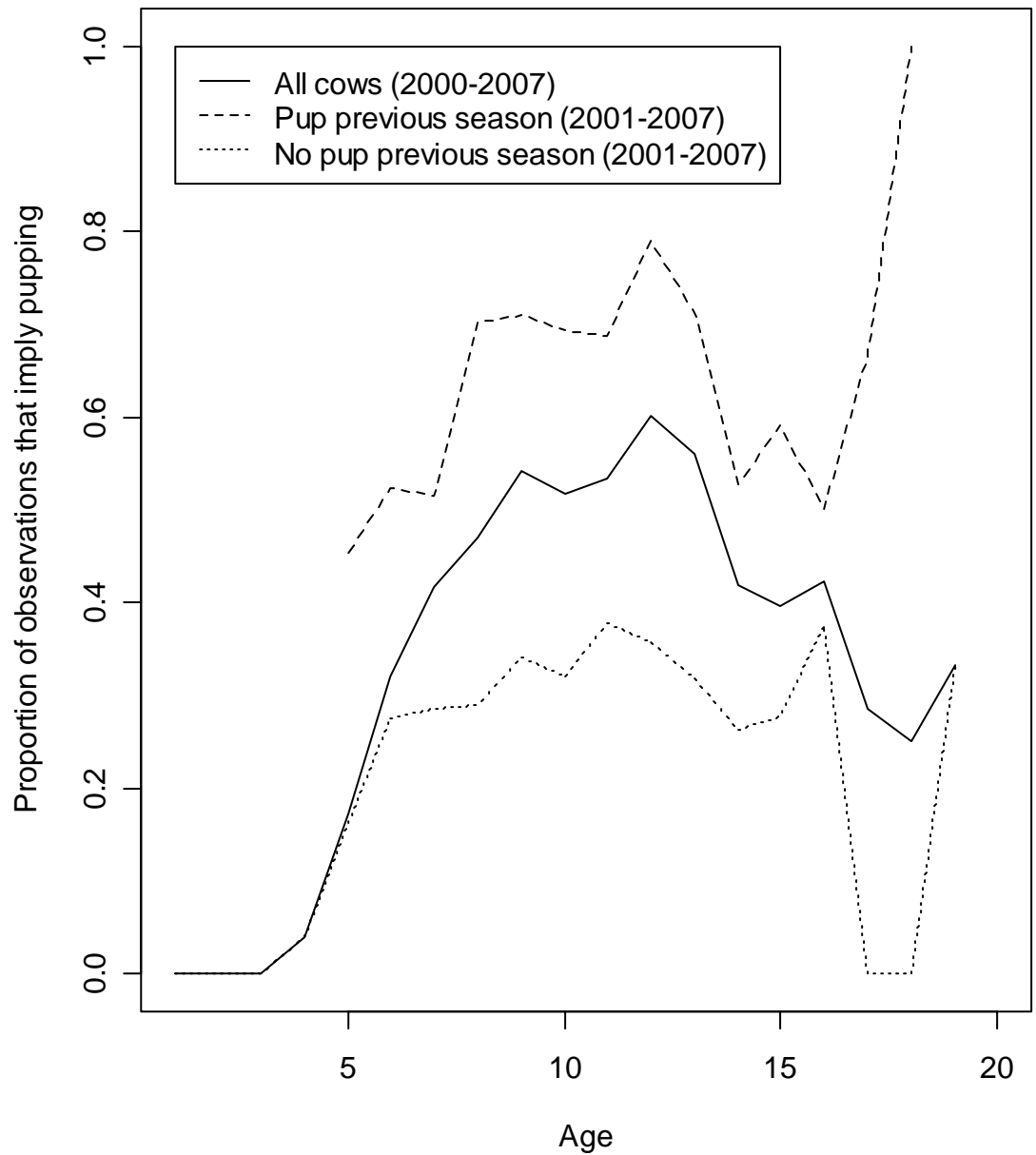


Figure 4. Pupping rate by age for tagged cows known to be alive based on Enderby Island observations after 1999. Figure shows mean for all cows, for cows that bore a pup in the previous season and for those that did not.