New Zealand Aquatic
Environment and Biodiversity
Report No. 15
2008
ISSN 1176-9440

Evaluating methods for estimating the incidental capture of New Zealand sea lions

Edward R. Abraham

Evaluating methods for estimating the incidental capture of New Zealand sea lions

Edward R. Abraham*

Datamine
P O Box 37120
Parnell
Auckland 1151

*Present address
Dragonfly, 10 Milne Terrace,
Island Bay
Wellington

New Zealand Aquatic Environment and Biodiversity Report No. 15
2008

Published by Ministry of Fisheries
Wellington
2008

ISSN 1176-9440

©
Ministry of Fisheries
2008

Citation:
Abraham, E.R. (2008).
Evaluating methods for estimating the incidental capture of New Zealand sea lions.
New Zealand Aquatic Environment and Biodiversity Report No. 15. 25 p.

This series continues the
Marine Biodiversity Biosecurity Report series
which ceased with No. 7 in February 2005.

EXECUTIVE SUMMARY

Abraham, E.R. (2007). Evaluating methods for estimating the incidental capture of New Zealand sea lions.

***New Zealand Aquatic Environment and Biodiversity Report No. 15.* 25 p.**

Methods for estimating the incidental catch of sea lions in the SQU 6T squid fishery are compared. A simulate and test technique is used, with a trial model being used to generate synthetic data that has the same statistical properties as the actual capture data. The estimation approaches are then tried on the simulated data. This has the advantages that the simulated captures are known exactly, and that large numbers of simulated data sets may be produced.

The trial model is fitted to observer data from the 2000 to 2004 calendar years. The simulated data are derived by applying the trial model to all the tows from the catch effort database that were within SQU 6T and which targeted squid, from the same time period. The trial model was a generalised linear model (GLM) with fixed effects and with year and vessel within year random effects. The tested estimation procedures were a fixed effects GLM, a GLM with both fixed and random effects, and a simple ratio estimator. The ratio estimate assumes that, within any year, the sea lion strike rate is constant across the fishery.

In order to reduce sea lion capture, sea lion exclusion devices (or SLEDs) were introduced into the fishery in 2000. Since 2004 they have been used on most tows. The use of SLEDs has reduced the number of sea lions seen by observers, and consequently the estimation problem has become more difficult. The performance of the estimators is compared under assumed levels of SLED coverage, ranging from 10% to 90%.

An assumption of the models used is that the SLED retention probability has remained constant throughout the period of the data. This assumption may not be a good one, as re-fitting the trial model to data from 2002 to 2004 produces a model with a higher retention probability and a lower estimate of the total sea lion interactions in years when SLED use was high.

The conclusion of the analysis is that the model based estimators are indistinguishable and they both consistently outperform the ratio estimator.

1 INTRODUCTION

New Zealand sea lions (*Phocarctos hookeri*) breed on the Auckland Islands. Their range overlaps with the Auckland Islands squid trawl fishery (Chilvers et al. 2005) and sea lions are caught in this fishery (Baird 2005a, 2005b, Smith & Baird 2005, 2006). In order to reduce the impact of the squid fishery on the sea lion population, a number of measures are used (Ministry of Fisheries 2006). Firstly, there is a geographical exclusion zone, with no trawling permitted within 12 nautical miles of the Auckland Islands. Secondly, a sea lion catch limit is set by the Ministry of Fisheries (known as the Fisheries Related Mortality Limit or FRML). If this incidental catch is calculated to have been exceeded during a season, then the squid fishery is closed. Thirdly, since 2000, sea lion exclusion devices (SLEDs) have been used on squid trawls. These are grates that fit across the entrance to the codend of the net, preventing sea lions from entering. An escape funnel in the net above the SLED allows sea lions to exit.

Given the economic impact of the regulation, the uncertainty of the effect of the fisheries-related mortality on the sea lion population (Breen & Kim 2006), and the controversy surrounding the sea lion catch (Forest and Bird 2006), it is important to be able to state what is known about the numbers of sea lions that are caught through fishing. The primary data on the incidental catch of sea lions are direct records by Ministry of Fisheries observers of sea lion capture. Typically, between 20% and 40% of the tows within the SQU 6T squid fishery are observed by Ministry of Fisheries observers, although in 2001 close to 100% of the tows were observed. The problem that is discussed in this report is how to extrapolate from these observations to a fishery-wide estimate of the sea lion captures. Since the introduction of SLEDs, the estimation problem has been made more difficult. Many of the animals that enter the nets now escape, so there are fewer direct observations on which to base the estimates. As it is not known what proportion of sea lions which leave the net through a SLED survive, mortality estimates are also more uncertain. We focus on estimates of the number of interactions between sea lions and trawl nets. This is the number of sea lions that enter a net, irrespective of whether they leave via a SLED or not, and irrespective of whether they are alive or dead when the net is brought on board. This provides an upper bound on the mortality, and if information on the survival rate of sea lions that have escaped a net through a SLED becomes available, may be used as a basis for a more precise mortality estimate.

Two main techniques for estimating interaction rates have been used. The first is a simple ratio method (Baird 2005a, 2005b). This uses the observer data to estimate an average strike rate, which is then applied to the non-observed trawls to obtain a fishery-wide total interaction. This ratio estimator relies on the assumption that the observer data are representative of the whole fishery. When SLEDs are used, a strike rate for trawls with and without SLEDs may be calculated and the two strike rates can then be appropriately applied to the non-observed trawls. A similar assumption is used in setting the FRML, as the method for calculating the limit assumes a constant sea lion strike rate across all tows. In a Bayesian model of the sea lion captures Breen et al. (2005) modelled the data, but assumed a single strike rate for each season. We compare the ratio method with approaches which assume that the strike rate may be spatially and temporally varying, within a season, and aim to predict it as a function of other factors.

The second technique uses a generalised linear model (GLM) to determine a relationship between the strike rate and factors such as the duration of a trawl, the distance of the trawl from the sea lion colonies, time of day, etc (Smith & Baird 2005, 2006). Given these factors, an estimate of the expected number of sea lions captured on a trawl can be derived. The modelling approach has the advantage that it is less prone to bias in the observations, and it allows estimates to be made in years when observer coverage is low. There is the danger, however, that models may over-fit the data, putting too much weight on the particular set of observations that are used to build them. In this case they may make unrealistic predictions when applied to the whole fishery data.

A difficulty with assessing the performance of the models is that the true interaction rates are unknown. In many modelling analyses, techniques such as using hold-out data or model cross-validation are used to gauge model performance. The problem here is that there are relatively few observations of sea lion capture (from a modelling point of view), and model performance may deteriorate rapidly if the model is fitted to a reduced data set.

In this report we discuss a method for overcoming this difficulty. We test the estimation methods on simulated data. Synthetic capture data are produced which have a similar statistical structure to the real data. These data are then used to test how well an estimation method built on an observed portion of a simulated set recovers the simulated captures. The first step is to fit a trial model to the observer data. This model is then used to generate a set of expected capture rates for all the trawls. From these expected rates, repeated sets of simulated captures are generated. A portion of each set is declared to be observed, and modelling procedures are checked to see how well they reproduce the actual synthetic captures (Figure 1).

We illustrate this procedure with three different candidate estimation approaches. Our intent is not to identify a definitive modelling method, but rather to illustrate how the simulate-and-test technique may be used to estimate model skill. We use data from 2000 through to 2004 (calendar years). This five year period covers the introduction of SLEDs to the squid fishery, and is a period for which estimates of sea lion capture have already been made.

2 METHODS

2.1 Data

2.1.1 Observer data

The key source of data on sea lion captures is the database of records returned by fisheries observers, *obs_lfs*[1]. This database contains a record for all non-fish bycatch recorded by the observers. The results of any autopsies are appended to the observer data, giving a confirmed species identification, sex, and other characteristics. For this project, an extract of the obs_lfs database was obtained, providing the core data for the analysis. All records were kept where the tow either started or finished within area SQU 6 or where either a sea lion or a fur seal had been caught. Records with fur seals were also retained as it was anticipated that sea lions would occasionally be mis-identified by observers as fur seals (and vice versa). Data were extracted covering the calendar years 2000 to 2004, inclusive. The data were restricted to trips with Ministry of Fisheries observers (these have trip numbers of less than 9000).

Within the time period concerned, 132 animals were identified by observers as sea lions. The identification recorded by the observer (*species_obs* records from the *obs_lfs* database) was checked against the identification recorded at autopsy (*species* records). Rates of misidentification were low. One animal identified by the observer as a fur seal was found to be a sea lion, one animal identified by the observer as a sea lion was found to be a fur seal, and one animal identified by the observer as a sea lion was found to be an elephant seal. There were 50 animals identified as sea lions by the observer which were not checked at autopsy. A single sea lion was recorded as decomposed and so assumed to have died before the trawl. We do not include this record as a sea lion capture in this analysis. After making these corrections, there were 130 sea lions recorded. Of these sea lions, seven captures were of live animals. Comments recorded by observers state that these animals in general appeared unharmed, and all were

[1] Documentation on this and other Ministry of Fisheries databases is available from their website, www.fish.govt.nz

returned to the sea. We include these animals as captures, however, as these events contain information on where and when sea lion captures are likely to occur.

Of these 130 sea lions, 11 were not caught in the squid fishery. Because capture rates outside the squid fishery are low, we exclude these animals and restrict our analysis to tows which were targeting squid. Of the remaining sea lions, seven were caught on the Stewart Snares shelf. In this region capture rates are too low to be reliably modelled and will be ignored. The geographic range of this study is restricted to the Auckland Islands part of the SQU 6T region, bounded by 51° 30'S, and 49° 30'S and by 165°E and 168°E. We are left with an observed catch of 112 sea lions over the five year period. For the purposes of estimating total sea lion catch, however, it must be born in mind that over this period a further 17% of this number were caught elsewhere or in other fisheries.

The *obs_lfs* database does not have information from all tows. In order to obtain positions and times of tows where sea lions were not caught, the data were merged with an extract from the obs database. All tows from trips that had data in the original *obs_lfs* extract, and which were within the study area, were selected. The spatial distribution of the sea lion captures and the observed tow positions are shown in Figures 2 and 3. Sea lion captures occur in two main regions, one on the shelf to the north of the Auckland Islands, and another to the east of the islands along the shelf edge. The captures on the Snares plateau can also be seen. This distribution is partly a reflection of the distribution of the positions of the observed trawls (Figure 3).

These data may be used to obtain an estimate of the probability that a sea lion is captured during a tow, as a function of tow position (Figure 4). The ratio of the numbers of tows where sea lions are caught to the number of observed tows is shown in the figure. This is a straightforward empirical estimate of the strike rate, and SLED use or other factors have not been allowed for. Because the capture rates are low, this ratio is calculated only where there are more than 20 observed tows. Strike rates are highest on the shelf region to the north of the Auckland Islands, reaching a peak of over five sea lions per 100 tows.

All times from the observer databases were converted from New Zealand Daylight Time to New Zealand Standard Time (UTC + 12 hours).

2.1.2 SLED data

From the 2000 season onward, sea lion exclusion devices were deployed in the squid fishery (Figure 6). Initially, they were used in selected tows to trial the methodology. From 2004 they have nominally been used on all tows within the squid fishery. The data on SLED use are not held in the Ministry of Fisheries databases. For this project, a spreadsheet prepared by Paul Starr of the Seafood Industry Council (SeaFIC) was used as the primary source. This covered SLED use on all observed tows during the 2000 to 2004 seasons. During the trial period, the SLEDs were deployed with a cover over the escape hole. Whether the SLED cover was open or closed is recorded. Following Smith & Baird (2005, 2006), we consider a net with a closed SLED to be equivalent to a net without a SLED. An open net has a SLED with an open escape; all other nets are closed. Through the remainder of the report, a net with a SLED is assumed to mean a net from which a sealion can escape, unless it is specifically indicated that the SLED is closed.

There are tows in our data which are not in the SeaFIC spreadsheet; many of these are from fisheries within SQU 6T which were not targeting squid. We assume that where SLED use is unknown, the net is closed.

2.1.3 Effort data

To allow extrapolation from the observed captures to an estimate of captures from all tows, data from the Ministry of Fisheries catch effort database are used. We extracted the time, position, and characteristics of all tows that started or ended in SQU 6T during the period of interest (January 2000 through to July 2004). The qualitative impression is that the distribution of the observed tows (Figure 3) reflects the distribution of fishing effort (Figure 5).

2.1.4 Data by year

A breakdown of the data by year is shown in Table 1. The total effort in SQU 6T varied widely, ranging between 582 and 2595 tows over the five year period. Similarly, observer coverage has varied, with close to 90% in 2001 and 27% coverage in 2003 and 2004. The number of open net tows has increased to 91% of all tows in 2004. Observed sea lion capture has varied between 11 and 39 sea lions in a year.

These data can be used to estimate an empirical strike rate (Table 2). Across the five year period the strike rate on closed net tows has been 0.059 (sea lions per tow). There is considerable inter-annual variability, with the strike rate reaching a low of 0.029 in 2003 to a high of 0.126 in 2001. The open net strike rate calculated from all the data is 0.018, and this has varied between 0 and 0.021 across the years. The SLED factor is the reduction in capture probability caused by using a SLED. Across all data, the SLEDs appear to reduce the capture rate to 30.1% of what it would otherwise be. Possibly because of the relatively low number of captures, there is considerable variability in this ratio when it is calculated for individual years. In Section 3.1.2 we discuss whether this retention probability can be assumed to be constant across all years.

Table 1: Summary of numbers of records used in this study, from the Auckland Islands part of the SQU 6T fishery, between 2000 and 2004.

	2000	2001	2002	2003	2004	Total
Total number of tows	1 206	582	1 646	1 466	2 595	7 495
Number of observed tows	437	529	562	397	719	2 644
Number of open-net observed tows	0	260	125	24	661	1 070
Number of observed sea lion captures	25	39	21	11	16	112

Table 2: Average strike rates and the efficiency of SLEDs. The table shows the strike rate (number of sea lions caught per tow) for each year, with and without open SLEDs. The third row gives an empirical estimate of the SLED efficiency, being the number of animals caught per tow when the net is open as a percentage of the number of animals caught per tow when the net is closed. The 2003 data for the open net strike rate and the SLED factor are excluded because of the low number of observations (24) of open net tows.

	2000	2001	2002	2003	2004	Average
Closed net	0.057	0.126	0.048	0.029	0.034	0.059
Open net	-	0.019	0.0	-	0.021	0.018
Empirical SLED factor	-	15.2%	0%	-	61.4%	30.1%

2.2 Trial model

2.2.1 Choice of model

The trial model is used to generate the simulated datasets. It is developed through analysis of the observer capture data. As a first step, it was assumed that the sea lion capture events are independent of one another. Previous modelling of the data using a Poisson distribution has found some evidence of over-dispersion (the actual number of observed multiple capture events is higher than would have been expected from the model) (Smith & Baird 2006). To allow for the possibility of some over-dispersion the sea lion capture is represented as a quasi-Poisson process. The distribution is not specified, but the relationship between the mean and the variance is assumed to be fixed, with the variance being larger than the mean by a factor which characterises the over-dispersion

We built a GLM that uses the observer data to estimate the dependence of the underlying rates on a number of known variables. Much of the model building effort concerns the preparation of plausible variables and their selection. Following Smith & Baird (2006), we include year and vessel-within-year random effects in the model.

2.2.2 Data preparation

A number of possible explanatory variables were prepared from the merged observer data. We followed the results of Smith & Baird (2005, 2006) in preparing these variables, but we also included additional positional and temporal variables to assess their possible significance.

SLED: A variable was introduced with a value of 0 if the net was closed and a value of 1 if it was open.

Tow duration: The logarithm of the duration of each tow was included in hours.

Daylight period: From the start and finish time of each tow we derived a set of factor variables which describe whether the tow spanned dawn or dusk, or was entirely at night or entirely during the day. The time of dawn or dusk was calculated using an astronomical algorithm (Meeus 1999), which estimated the time when the sun would be 6° below the horizon at the position and date of the tow.2

Time of day: To allow for further dependence on time of day, a sinusoidal function of the hour of the day was introduced. Two variables, $chour = \cos(2\pi\, hour/24)$ and $shour = \sin(2\pi\, hour/24)$ were defined, where *hour* is the start time of the tow.

Time of year: Similarly, a sinusoidal function of the year day was introduced. Two variables, $cyday = \cos(2\pi\, year\_day/365.25)$ and $syday = \sin(2\pi\, year\_day/365.25)$ were defined, where *year_day* is the Julian day.

Position: The start position of the tow was included (latitude and longitude, with latitude south being negative). An offset was subtracted from the positions (-50 from the latitude and 166 from the longitude), to reduce the effect of these terms on the model intercept.

Using code provided by Murray Smith (Smith & Baird 2005, 2006), the distance of the sea lions from each of the five main colonies on the Auckland Islands was calculated. These colonies and their latitudes and longitudes are Dundas (-50.578, 166.319), Sandy Bay (-50.502, 166.281), Southeast point (-50.504,

[2] Available from http://www.sci.fi/~benefon/rscalc_cpp.html (December 2006)

166.325), Figure of Eight west arm (-50.847, 165.899), and Figure of Eight east arm, (-50.870, 166.218). The great circle distance (in km) was calculated. No adjustment was made where the shortest path crossed over land. The positions given for the Figure of Eight colony are two points on either side of the harbour entrance. An extra 20 km was added to account for travel from the colony to the harbour mouth. The logarithm of each of these distances was trialled in the model, as well as the logarithm of a variable 'colony distance' which was the minimum of those values.

Following the observation that much of the capture occurs close to the shelf edge, we also tested a variable that gave the distance to the shelf edge (defined as the 400 m depth contour). This variable was not preferred by the models.

The captures occur in two main areas within SQU 6T. To allow for the possibility that capture rates differ between these regions, a factor was introduced that identified in which of these areas a tow was made. If the longitude/2 + latitude was greater than 133.83, or the longitude was greater than 166.94, then the area factor had a value of 'East'. Otherwise it had a value of 'North'.

Fishing type: Factors were included that indicated whether the trawl type was bottom ('BT') or mid-water ('MW').

2.2.3 Model selection

We modelled the observer data through a GLM with random effects. The appropriate link function for a quasi-Poisson model is the log function. This means that the resulting model is multiplicative. Fitting was carried out within R (version 2.4.0, R Development Core Team 2006), using the 'lmer' function from the 'lme4' package (version 0.9975-10, Bates & Sarkar 2006), as this function allows the fitting of random effects terms within GLMs. The random effects were included at the year and vessel within year levels. The Laplace method was used for carrying out the model fits.

A greedy stepping procedure was used to determine which terms to include in the trial model. Firstly, all potential terms were tried individually in the model, and the term which caused the biggest decrease in Mallows C_p (Mallows 1973) was retained. This is evaluated within the 'lmer' function as the deviance plus twice the number of degrees of freedom in the model terms. The Akaike Information Criterion, or AIC, is often used in model selection; however it cannot be calculated for the quasi-likelihood models (Venables & Ripley 2002), and so C_p is used instead. Confusingly, C_p is referred to within 'lmer' as the AIC.

The procedure was then repeated, sequentially adding one of the remaining terms. We stopped adding terms when either no terms caused C_p to decrease, or when the number of terms in the model reached a preset maximum. This arbitrary rule ensured that the model building could be run as a fully automated procedure, without allowing the inclusion of a large number of terms which have a relatively minor effect on the model predictions. In building the trial model, the maximum number of terms was set to nine, otherwise the number of terms was set at five. The five term model which was fitted to the original data is referred to as the best model. The over-fitting of the trial model relative to the best model means that the trial model had variation which was unable to be captured by the test models, but which had a correlation that was more realistic than random noise.

2.3 Simulation

The resulting model trial model was then used to generate a set of Poisson means for each tow in the catch-effort database that was within SQU 6T and which targeted squid. The Poisson mean is the model estimate of the average expected catch on that tow. In generating the model estimate, it is assumed that the net is closed. From the Poisson means, simulated data are then produced by sampling from the Poisson distribution. These data reflect a simulated sea lion net interaction.

In order to generate the simulated data, it was necessary to assign an observer and a SLED status to each tow. An observer coverage of 30% was assumed. Actual observer status of the trips was not determined, rather 30% of the trips were randomly assigned observers, and all tows on those trips were taken to be observed. Simulations were made with various levels of SLED usage (either 10%, 30%, 50%, 70%, or 90%). For example, to achieve a required SLED coverage of 50%, all tows on a randomly chosen 50% of trips were assumed to have been carried out with SLEDs. The exact proportion was then obtained by randomly flipping the SLED status on individual tows. This scheme meant that SLED use tended to be concentrated within individual trips. For both the simulated SLED usage and the observer coverage, all the data were taken together, and no attempt was made to achieve exactly the same proportion within each year. Where SLEDs were used, the number of captures on each tow is set to zero with a probability of 0.8. This reflects an assumed escape probability of 20%. For each level of SLED usage, 40 simulated sets were made, each set having a simulated capture associated with every tow in the catch-effort data.

2.3.1 The tested procedures

The following three estimation methods were then tested on the synthetic data:

1) Ratio estimation. This is direct estimation of the total interaction using known observer coverage and a known SLED factor. If the SLED retention probability is s, the number of tows which were observed is t_{obs}, the total number of tows is t_{total}, the number of sea lions which were caught on observed tows with open SLEDs is n_{SLED}, and the number of sea lions which were caught on observed tows without SLEDs is n_{closed}, then the ratio estimate of the total number of sea lion interactions is $n_{sealions} = (n_{closed} + n_{SLED}/s)\ t_{total}/t_{obs}$

2) Fixed effects GLM. A quasi-Poisson distribution is assumed, with a logarithmic link function. There are no random effects terms, and the year of each tow is included as a factor amongst the covariates. The same model fitting procedure is used as for the best model, with a maximum of five terms being included in the model.

3) Random effects GLM with year and vessel within year random effects. This is the same model structure as best model and the trial model. The number of terms which are included in the models is restricted to a maximum of five.

3 RESULTS

3.1 The trial model

The initial fit to the data was carried out with a mixed-effects GLM. The full trial model, with nine fixed effects terms, is shown in Table 3. The variable with the greatest explanatory power is whether or not an open SLED is used. The most significant spatial variable is the longitude, with a higher catch rate being predicted in the east of the Auckland Islands part of SQU 6. The reduction in the model deviance is less than 1% for the remaining terms (Table 4). If only the first five terms are retained, then we obtain a candidate for the best model (Table 5). In addition to SLED use and longitude, this model has terms that relate to the tow duration, the distance from Sandy Bay, and the cosine of the time of day. The dependence of the model on the duration of the tow is sub-linear, being close to a square root.

Table 3: Coefficients and confidence intervals of covariates in the trial model of the observer data, with nine fixed effects retained in the model. The coefficients are multiplicative. The year random effects have a standard deviation of 0.62 and the vessel within year random effects have a standard deviation of 0.33. The dispersion of the model is $\sigma = 1.022$.

Covariate	Value (95% c.i.)	Significance
Intercept	0.3 (-2.46, 3.06)	
SLED	-1.71 (-2.4, -1.02)	**
Longitude – 166	-1.03 (-2.1, 0.03)	.
log(Tow duration)	0.39 (-0.07, 0.85)	.
log(Distance to Sandy Bay)	-3.36 (-6.16, -0.57)	*
cos(Hour×π/180)	0.35 (0.02, 0.68)	*
log(Colony distance)	2.61 (-0.25, 5.47)	.
sin(Hour×π/180)	0.52 (0.13, 0.9)	**
Dusk tow	0.77 (0.06, 1.48)	*
cos(Year day×π/180)	-0.62 (-1.34, 0.09)	.

Table 4: Summary of the change in deviance in the trial model as terms are sequentially added. Terms which produce the biggest decrease in the C_p criterion are included first.

Covariate	Deviance	% Reduction	C_p
Intercept	738.7		
Random effects	719.2	2.6	725.2
SLED	695.9	3.2	703.9
Longitude – 166	675.6	2.9	685.6
log(Tow duration)	669.7	0.9	681.7
log(Distance to Sandy Bay)	664.6	0.8	678.6
cos(Hour×π/180)	660.7	0.6	676.7
log(Colony distance)	657.7	0.5	675.7
sin(Hour×π/180)	654.3	0.5	674.3
Dusk tow	650.2	0.6	672.2
cos(Year day×π/180)	647.5	0.4	671.5

Table 5: Coefficients and confidence intervals of covariates in the 'best' model of the observer data. Only five fixed effects are retained in the model. The coefficients are multiplicative. The year random effects have a standard deviation of 0.49 and the vessel within year random effects have a standard deviation of 0.36. The dispersion of the model is $\sigma = 1.023$.

Covariate	Value (95% c.i.)	Significance
Intercept	1 (-1.71, 3.71)	
SLED	-1.74 (-2.41, -1.06)	**
Longitude – 166	-1.5 (-2.52, -0.49)	**
log(Tow duration)	0.53 (0.09, 0.96)	*
log(Distance to Sandy Bay)	-0.96 (-1.7, -0.22)	*
cos(Hour×π/180)	0.29 (0.00, 0.59)	.

The best model is similar to the one derived by Smith & Baird (2006). In their model, the fixed effects are the log of the minimum colony distance, a factor which classifies the tow by time of day relative to dawn or dusk, a SLED factor, and the log of the tow duration. As here, the random effects are year and vessel. They find that a net with an open SLED reduces the catch rate by 0.27 ± 0.10 and that the catch rate scales as the tow duration raised to the power of 0.58 ± 0.17. In comparison, we find that an open SLED reduces the catch rate by a factor of 0.18 (95% c.i. of 0.08 to 0.35), and that the sea lion capture rate scales as the tow duration raised to the power of 0.53 ± 0.43. It is not surprising that different model terms come to the fore. The Smith & Baird (2006) model was built on data spanning a longer period, and because of the few capture observations the term selection may be sensitive to the particular observations used. They also included different spatial terms, with the longitude covariate being unavailable to their model.

In order to assess the fit of the model, the relationship between the predictions from the best model and the actual capture data is shown in Figure 7. The data are first smoothed with a Friedman super smoother, with a span of 0.1 (Friedman (1984), available within R as the function 'supsmu'). A linear relationship is then seen, which closely follows the 1:1 line. A comparison between the predicted capture rates of the best and the trial models is shown when they are applied to the catch-effort data. The standard deviation of the difference between the logarithm of the predicted capture rates is 0.47. The detail of the trial model cannot be captured with the smaller number of terms permitted for either the best model or the test models.

3.1.1 Predicted sea lion captures

The predicted sea lion capture rates are compared with the observed rates in Figure 9. The empirical capture rate for each year is shown compared with the best model predictions from the observer data. The model fits the data well, although this is unsurprising as the inclusion of a year random effect gives it considerable freedom to fit the annual captures. The predicted total capture rate (sea lions which enter the net, irrespective of whether they escape via a SLED) is also shown, based on the observer data and on the catch-effort data. There is close agreement between predictions from the observer and from the effort data, suggesting that the observer data are not providing a biased sample of the total fishing effort. The peaks in the total interactions in 2001 and 2004 reflect the high use of SLEDs. In these years many sea lions would have escaped from the net.

The spatial distribution of the predicted total interactions (from the trial model) is shown in Figure 10. This is the sum of the expected number of interactions for all tows from the catch effort data (see Figure 5), shown within 0.2° squares.

The total number of interactions which are predicted from the catch effort data are also shown in Figure 11. Results obtained using four different estimators are shown, the ratio estimate, and the three different models. The data are also given in Table 6. The models are all in close agreement with one another, with the exception of some spread in 2004. The ratio estimate is higher in 2001 and 2002 than expected from the models. The high values in 2004 of between 227 and 295 sea lion interactions stand out. The 2004 data are associated with high SLED use. Despite the large number, this is within the range estimated by Smith & Baird (2006) for 2004. They derive a mean value of 185 interactions with a quartile range of 134 to 219 sea lions and an upper 97.5 percentile of 376 interactions. In the next section we discuss why the 2004 value may be so high.

Table 6: Estimates of the total interactions (the number of sea lions that would have been caught if there were no SLEDs used). Four different estimators are used, the ratio method, the fixed effects model, the mixed effects model with a restricted number of terms, and the mixed effects model with a large number of terms which was used as the trial model. The data are shown graphically in Figure 11.

Estimator	2000	2001	2002	2003	2004
Ratio	69	127.3	116.2	50.4	270.1
Fixed effects	70.9	72.9	71.7	36.4	295.6
Mixed effects - best model	64.2	64.3	77	51	227.3
Mixed effects - trial model	59.4	60.7	69.8	44.8	278.7

3.1.2 The 2004 peak

The discrepancy between the observed captures and the predicted total interactions in 2004 is puzzling. There was close to complete SLED use reported for 2004, but the number of sea lions observed to be caught was similar to that in 2003. One reason for the increase is that there were considerably more tows in 2004 than in 2003 (an increase in number of tows of over 75%). In addition, some of the fixed effects show an increase over this period. For example, the tow duration has increased steadily since 2001, with an increase of over 50 minutes (20%) in the average tow duration seen in the effort data between 2003 and 2004 (Figure 12). This change is not due to outliers, rather there has been a consistent change in the distribution of the tow duration (Figure 13).

The year random effects are shown in Figure 14. Although the sea lion catch was unusually high in 2001, the random effects are all close to one for the other years. There is no sign from the random effects that the model had difficulty explaining the 2004 data.

The best model suggests that the SLED retention probability is 18% (with a 95% c.i. of 9% to 37%). Although there is a wide range in the confidence interval, the expected value is lower than has been found by others. Smith & Baird (2006) found the median posterior SLED retention probability is 26.5% (2.5% and 97.5% percentiles of 11.3% and 50%). Fitting a model to data from 2002 to 2005, Breen et al. (2005) find the median posterior SLED retention probability is 30.5% (5% and 95% percentiles of 16.6% and 56.4%). In addition, a direct estimate of the SLED retention probability from the observer data is 30.1% (see Table 2).

When SLEDs are widely used, as happened in 2004, the observed catches must be multiplied by a factor which accounts for the retention probability. If this probability is low, then the multiplier will be correspondingly high. An assumption of the modelling is that the SLED retention probability has remained constant from 2001 to 2004. This is the same assumption which was made by Smith & Baird (2006). Breen et al. (2005) explored whether the data suggest that the retention probability has changed, but found that allowing an annually varying retention probability made little difference to their predictions. However, they were only using data from the 2002 season onwards. As a test, we rebuilt the best model, restricting the data to the period from 2002 to 2004. The SLED retention probability increases to 35% (95% c.i. of 16% to 74%), and the expected total number of captures in 2004 falls to 108. The assumption that the SLED retention probability has remained constant has a marked effect on the predictions, and may be a poor one. The design has changed over the years. In particular, the escape hood now faces forward, so that sea lions have to actively swim to escape. The bar spacing has also been standardised. Unfortunately, information on the individual SLEDs is not available, only whether they were deployed open or closed, so there is no additional data that can be used to help determine the retention probability.

For this report, we continue with a set of models that assume a constant retention probability.

3.2 Performance of the estimators

3.2.1 Relative error

The ability of the models to recover the test data is shown in Figure 15, which shows the standard error in the difference between the predicted interactions and the actual interactions, normalised by the square root of the number of actual interactions. With this normalisation the error is similar between years (for the same level of SLED usage), even though the number of simulated captures in any year varies from 22 to 340. To calculate the standard deviation, the data are grouped by year and by SLED usage. The error bars in the figure show the range of the measure across the five years, for each value of the assumed SLED usage. The error measure may be converted to a coefficient of variation (c.v.) by dividing by the square root of the number of interactions.

There is no evidence of bias in the mean prediction errors. First finding the mean prediction error, averaged within each level of SLED usage, and then taking the maximum absolute value of these means, the mixed effects model has a maximum average error of 0.084, the fixed effects model of 0.14, and the ratio method of 0.43. Although the mean error in the ratio method is relatively large, it is not consistently biased high or low.

The two modelling methods are indistinguishable, and both are more accurate than the ratio estimator. This is even though the conditions are ideal for making the ratio estimate, as observers were strictly randomly assigned to trips and the SLED factor was known exactly. In the models, the SLED factor was derived as part of the model fit. There is a steady decrease in model skill as the SLED coverage increases.

3.2.2 Recovering the SLED retention factor

Because the process which was used to generate the data is known, the simulation can be used to test how accurately the parameters of the trial model can be recovered. As an example, the ability of the model to recover the SLED retention probability is shown in Figure 16. The skill of this estimate remains steady through much of the range, with an increase in the uncertainty as the SLED coverage decreases to 10%.

3.2.3 Caveats

Because the simulated data are generated by a known process, the only noise that is introduced is due to the sampling of the Poisson distribution. In reality, the variables controlling the probability a sea lion will enter a net are unknown. Whatever they are, they won't correspond to the simple expressions used here. This was represented by using an over fitted model as the trial model, but the models had the advantage that exactly the same terms were available for fitting the data as were used for generating it. For this reason, the true error in model estimates of sea lion captures is likely to be larger than we have determined.

4 ISSUES AND RECOMMENDATIONS

4.1 Data on SLED use

Accurate information on SLED use during observed tows is essential (including tows where sea lions are not caught). Currently the observers do not record SLED use directly, and so the data are not available from the Ministry of Fisheries records. Rather, SLED use must be derived from records compiled by SeaFIC. This greatly increases the potential for misalignment between the observer, effort, and SLED data. Currently, the SLED data are available only in informal spreadsheets. Given the value of these data, it should be incorporated into the Ministry of Fisheries database. The SLED design has changed with time, and may continue to change. Some more detailed information on the SLEDs should also be recorded, at least once per trip. This may allow a better assessment of how the retention probability is varying between trips and years.

4.2 Linking effort and observer data

We abandoned attempts to link effort and observer data. There are too many inconsistencies in times and positions to allow a simple reconciliation. On some trips there are more tows observed than there are tows in the effort database, and on other trips there are fewer. To allow better estimation it is essential that the observer and effort data can be reliably linked. This would provide a valuable integrity check on the datasets. This work is best done by the Ministry of Fisheries so a consistent linkage between the observer and effort data is available for all users. Ideally, tow identification numbers should be recorded by observers to allow the two sets of records to be directly related to one another.

4.3 The best estimator

Model estimators of sea lion captures outperformed the simple ratio estimator, at all levels of SLED use. There was little difference in the performance of the GLMs fitted with either fixed or mixed effects. The fixed effects GLMs were, however, simpler to implement. It would be interesting to compare the results presented here with those derived using the Bayesian methods of Smith & Baird (2006). Unfortunately, the Bayesian methods are computationally intensive and fitting the large number of test data sets using a Bayesian approach would be laborious.

4.4 The SLED retention probability

When SLED use is high, an accurate estimate of the SLED retention probability becomes important. It appears that the assumption of constant SLED retention probability may be a poor one. The unusually high values for the 2004 sea lion capture were greatly reduced when only a subset of the data was fitted. This suggests that the retention probability has changed.

5 SUMMARY

The procedure presented here, of building a simulated dataset against which to test candidate models, provides a way of directly comparing the skill of the various methods. Rather than relying on theoretical arguments that may or may not hold with the particular datasets encountered in the analysis of sea lion captures, the model approaches are trialled on realistic but controlled data. The result is clear, that a model based analysis performs better than the ratio estimate. We will make the data needed to repeat the simulations carried out here available via the Ministry of Fisheries. The intent is to provide a common ground that can be used for evaluating models.

There are other problems that could be looked at with this dataset. We focused on the issue of SLED usage and how that will affect the accuracy of estimates of the total sea lion interactions. The same method could be used to explore the consequences of changing the level of observer coverage.

6 ACKNOWLEDGEMENTS

We are grateful to Susan Waugh of the Ministry of Fisheries for her guidance on this project. Murray Smith of NIWA helped with the collation of the data and provided samples of code that he had previously developed. His assistance is much appreciated. Thanks are also due to Kim George of the Ministry of Fisheries for providing the data, and to Brian Bull and Geoff Grimwood of Datamine for suggestions on the statistical methods. The work was initially presented to the Aquatic Environment Working Group, and many useful comments and suggestions were received from that forum. This report was completed for the Ministry of Fisheries as part of ENV2005/02.

7 REFERENCES

Baird, S.J. (2005a). Incidental capture of *Phocarctos hookeri* (New Zealand sea lions) in New Zealand commercial fisheries, 2001–02. *New Zealand Fisheries Assessment Report 2005/8*. 17 p.

Baird, S.J. (2005b). Incidental capture of *Phocarctos hookeri* (New Zealand sea lions) in New Zealand commercial fisheries, 2002–03. *New Zealand Fisheries Assessment Report 2005/9*. 13 p.

Bates, D.; Sarkar, D. (2006). Linear mixed-effects models using S4 classes, version 0.9975-10. Retrieved November, 2006, from http://cran.r-project.org/doc/packages/lme4.pdf

Breen, P.; Kim, S.; Starr, P. (2005). *Estimating sea lion bycatch in the SQU 6T and SQU 1T fisheries*. NIWA Client Report: WLG2005-37, prepared for the Squid Fishery Management Company (Wellington: NIWA).

Breen, P.A.; Kim, S.W. (2006). Exploring alternative management procedures for controlling bycatch of Hooker's sea lions in the SQU 6T arrow squid fishery. Final Research Report, 88 p. (Unpublished report held by the Ministry of Fisheries, Wellington).

Chilvers, B.L.; Wilkinson, I.S.; Duignan, P.J.; Gemmell, N.J. (2005). Summer foraging areas for lactating New Zealand sea lions *Phocarctos hookeri*. *Marine Ecology Progress Series 304*: 235–247.

Forest and Bird (2006). Forest and Bird submission on the 2006-07 SQU 6T sea lion operational plan (Wellington: Forest and Bird). Retrieved December 2006, from http://www.forestandbird.org.nz/Submissions/submission_sealionssquid.pdf

Friedman, J.H. (1984). A variable span scatterplot smoother. Laboratory for Computational Statistics, Stanford University Technical Report No. 5. Stanford University, Stanford, CA.

Mallows, C.L. (1973). Some comments on C_p. *Technometrics 15*: 661–675.

Meeus, J. (1999). Astronomical algorithms (2nd edition). Willmann-Bell, Richmond, VA.

Ministry of Fisheries (2006). 2006-07 SQU 6T Sea lion operational plan, 14 September 2006. (Wellington: Ministry of Fisheries). Retrieved December 2006, from http://www.fish.govt.nz/NR/rdonlyres/8AE21896-401E-4AEC-BA69-AAE3B9923E1E/2847/200607FinalSQU6TIPP14sept1.pdf

R Development Core Team. (2006). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna.

Smith, M.H.; Baird, S.J. (2005). Factors that may influence the level of incidental mortality of New Zealand sea lions (*Phocarctos hookeri*) in the squid (Notodarus spp.) trawl fishery in SQU 6T. *New Zealand Fisheries Assessment Report 2005/20*. 35 p.

Smith, M.H.; Baird, S.J. (2006). Estimation of incidental captures of New Zealand sea lions (*Phocarctos hookeri*) in New Zealand fisheries in 2003-04, with particular reference to the SQU 6T squid (Notodarus spp.) trawl fishery. Aquatic Environment Working Group paper 2006/21, 33 p. (Unpublished report held by the Ministry of Fisheries, Wellington).

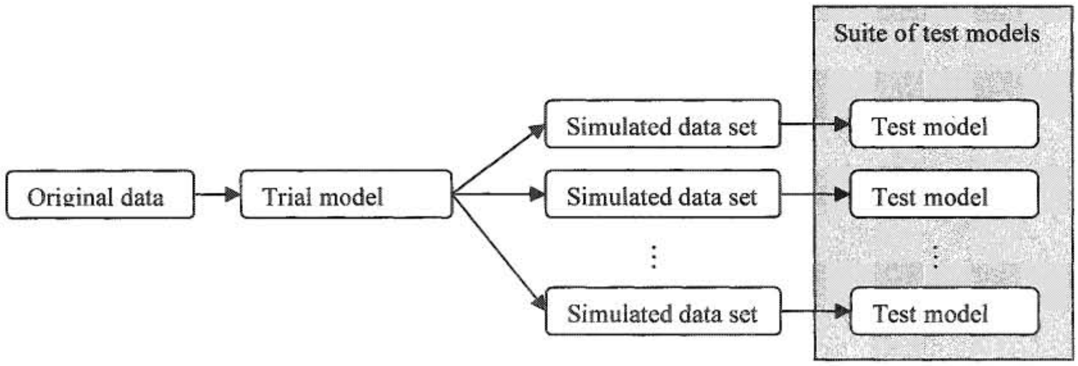

Figure 1: Schematic diagram of the procedure used in this report. A statistical model is fitted to the original data set (the trial model) and this is used to generate simulated data sets with a similar statistical structure to the original data set. Model selection and fitting procedures are then tested on the simulated data. Because the number of sea lion strikes in each simulated data set is known exactly, the model skill can be accurately evaluated.

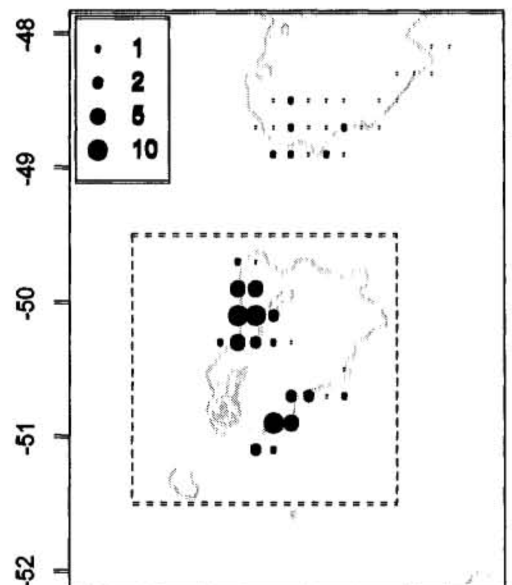

Figure 2: The number of sea lions recorded by observers within 0.2° squares, between January 2000 and July 2004. Squares where more than 20 observed tows were made, but where no sea lions were caught, are shown by dots. The map shows the Auckland Islands, the 200 m depth contour and the boundary of the Auckland Islands part of the SQU 6 fishing area.

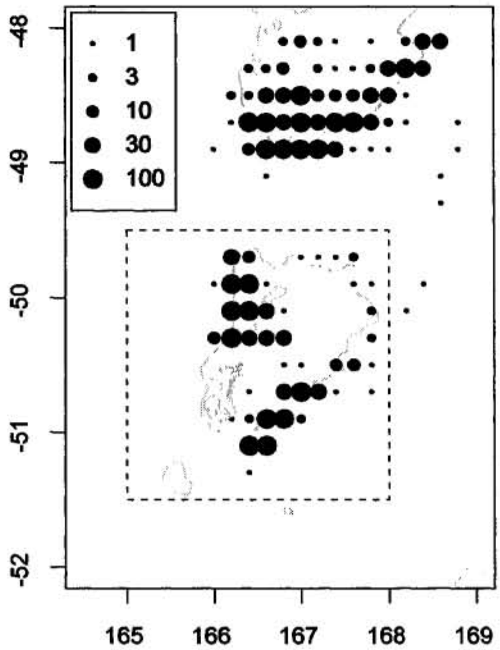

Figure 3: Numbers of observed trawl tows that targeted squid, within 0.2° squares, between January 2000 and July 2004.

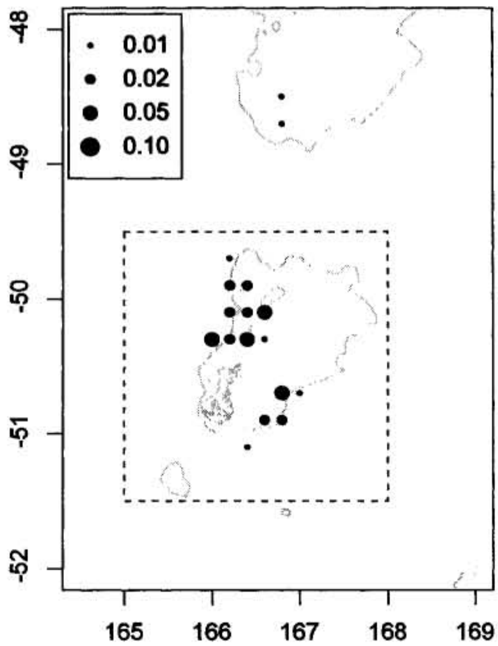

Figure 4: Empirical estimate of the strike rate (sealions per tow), as a function of location. This is the ratio of observed sea lion fatalities to the number of observed tows, within 0.2 ° squares. Only squares with 20 or more observed tows are shown.

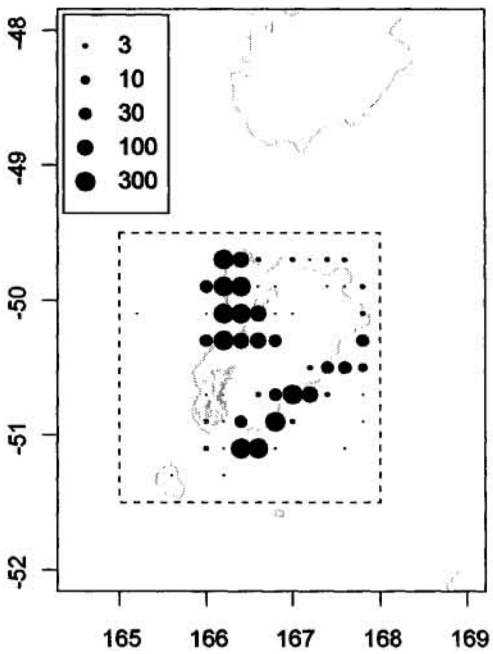

Figure 5: Numbers of tows between January 2000 and July 2004 from the catch effort database that targeted squid, within the Auckland Islands part of SQU 6. The number of tows is shown within 0.2 ° squares.

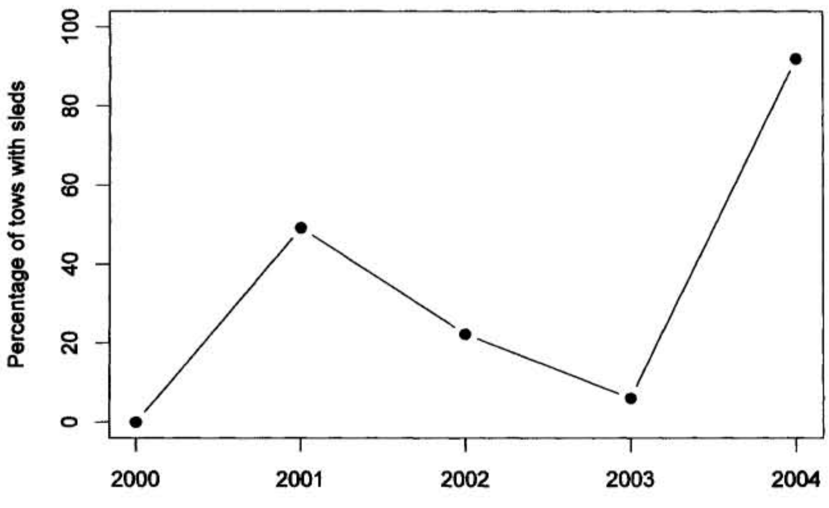

Figure 6: Percentage of tows which use open SLEDs, from SeaFIC data. During initial trials of the SLEDs in 2000, all the cover nets were closed and there were no open SLEDs. In 2004 over 90% of tows used open SLEDs.

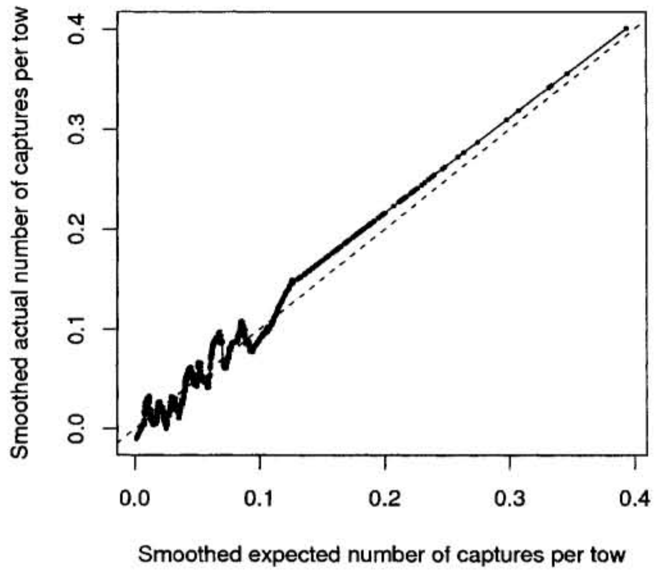

Figure 7: Smoothed model predictions. The actual captures are compared with the predictions from the best model, with the data being jointly smoothed. The 1:1 relationship is shown by the dashed line.

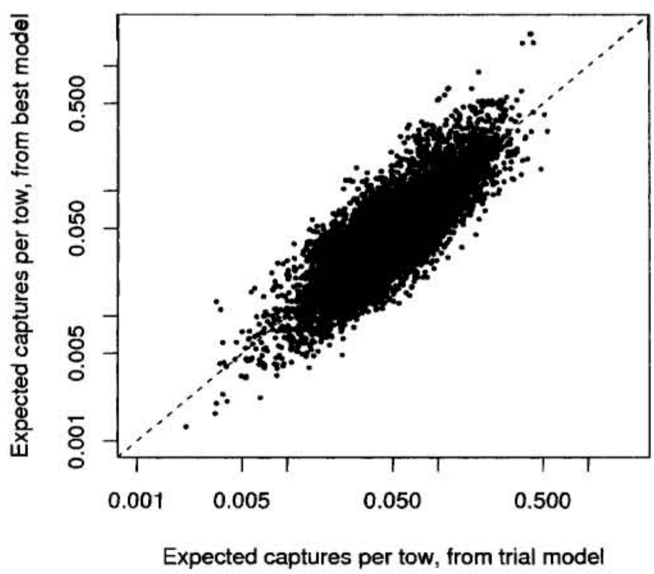

Figure 8: Comparison between the best and the trial models. The predictions of the best and the trial models, from the catch-effort data, are shown against one another.

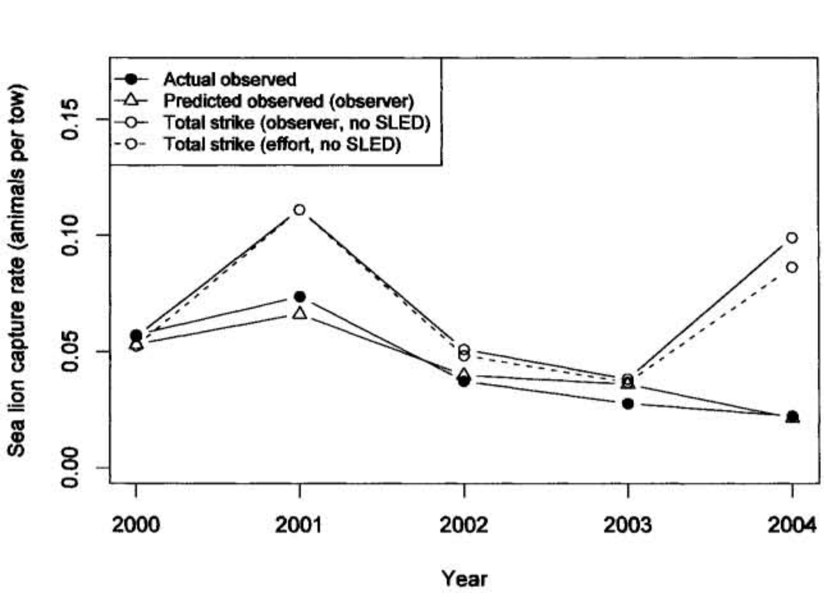

Figure 9: The sea lion capture rate per tow. The actual capture rate from the observer data is shown together with the model predicted observed capture rate (based on the observer data). In addition, the mixed-effects best-model predictions for the total strike rate are shown on the assumption that all nets were closed, with no open SLEDs being used. The predictions are shown for the strike rate based on all tows (the catch effort data) and on the observed tows.

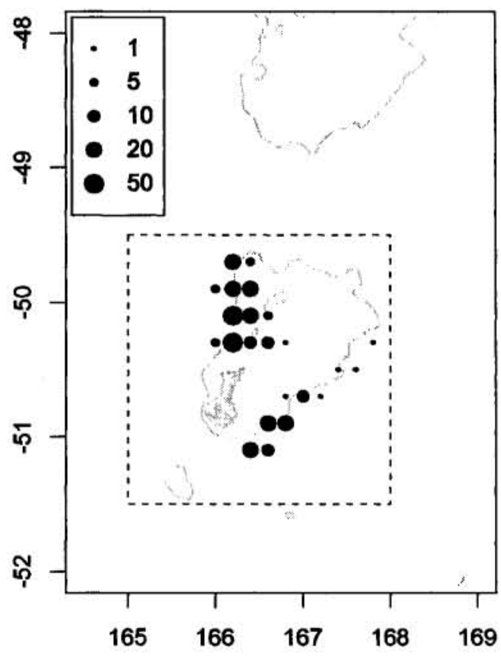

Figure 10: The spatial distribution of the predicted total number of interactions, over the five years of the data. The predictions are made by applying the trial model to the catch effort data shown in Figure 5.

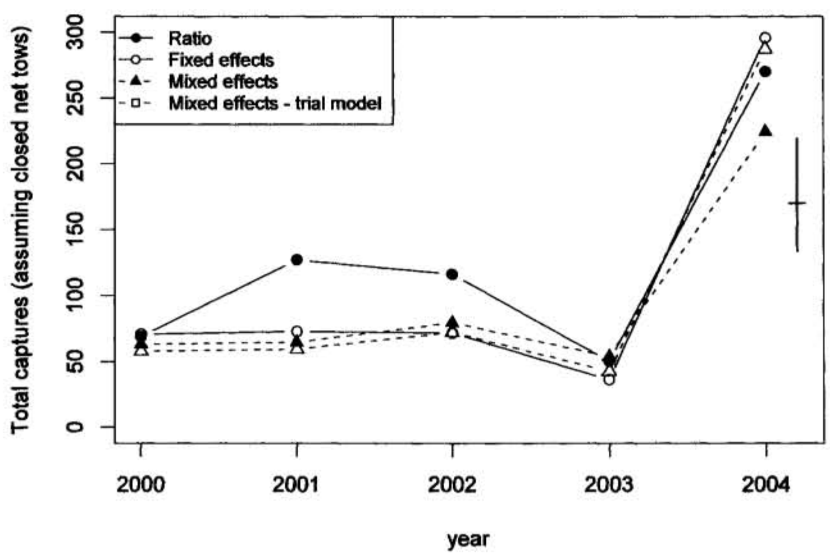

Figure 11: Predicted total sea lion strikes. This is the total number of captures which is predicted to have occurred if SLEDs were not used on the tows. Four different estimators are used, the simple ratio estimate, the fixed-effects GLM, the mixed-effects GLM, and the trial model, which has more terms than would be accepted in the usual model fitting procedure. The line to the far right shows the 25th, 50th, and 75th percentiles of the estimated total interactions in 2004, from the Bayesian model of Smith & Baird (2006).

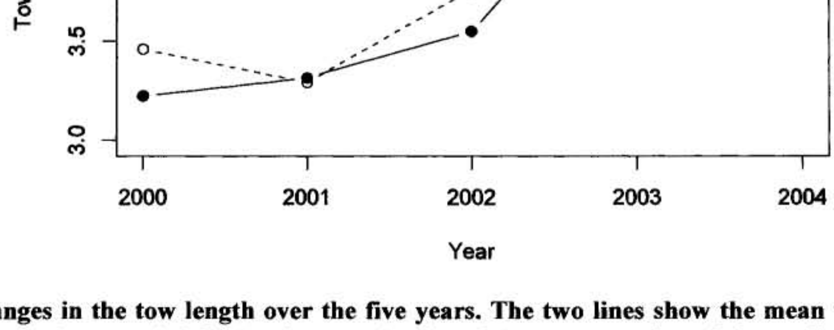

Figure 12: Changes in the tow length over the five years. The two lines show the mean tow length from the observer tows only (as recorded in the observer data) and the mean tow length from all tows (as recorded in the catch effort data). Both datasets show an increase in the mean tow length over the five year period.

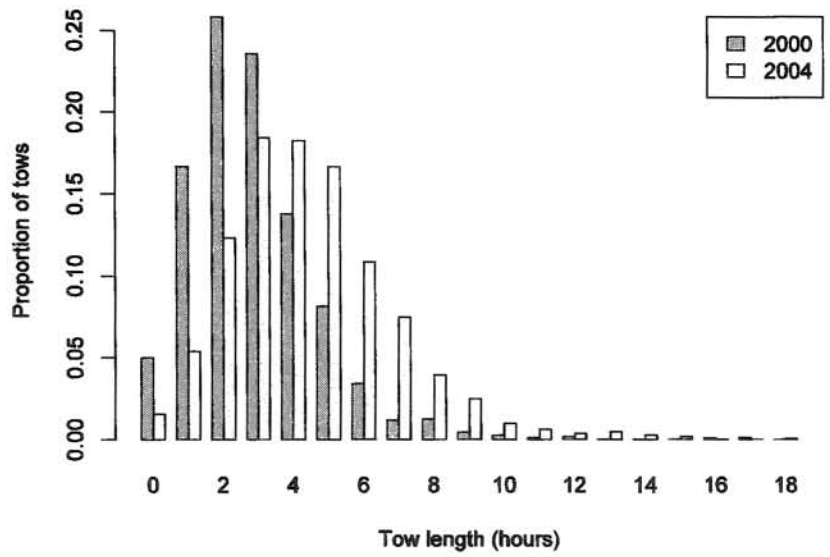

Figure 13: Change in distribution of tow length between 2000 and 2004, from the catch effort data.

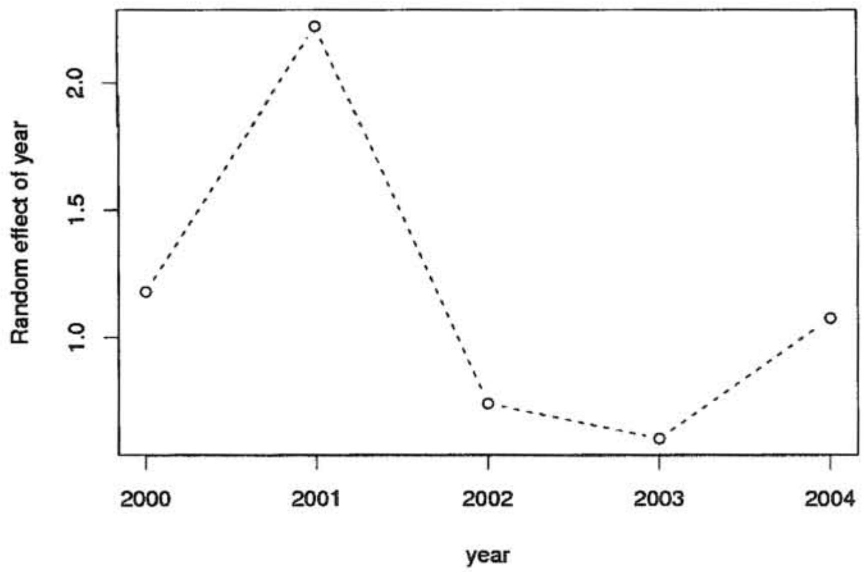

Figure 14: The year random effects. Catch rates are predicted to have been higher in 2001 than can be explained by the fixed effects alone.

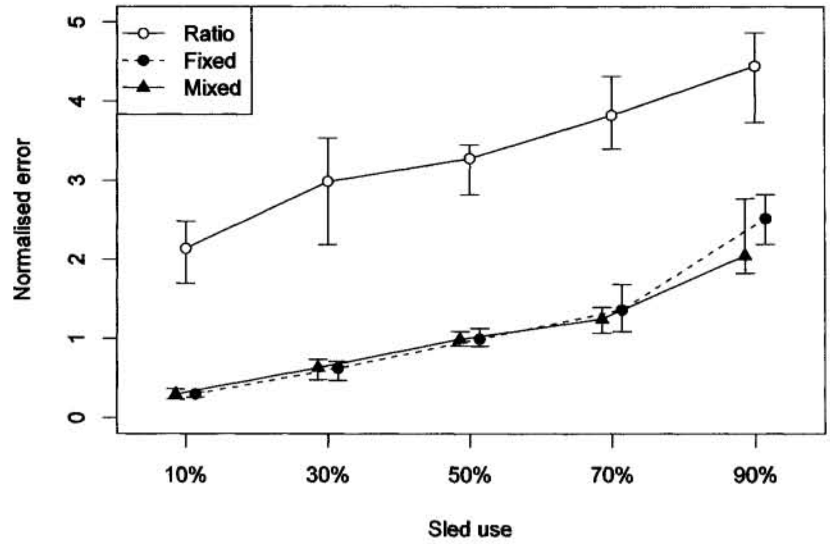

Figure 15: Comparing the models' ability to recover the test data. The test data are a simulation of the 2000 to 2004 data, with varying levels of SLED use. The y-axis shows the mean and range of the normalised error (predicted total – actual total)/sqrt(actual total), where the data are grouped by SLED coverage and by year. The three lines are for the ratio estimate, and the models with fixed and with mixed effects.

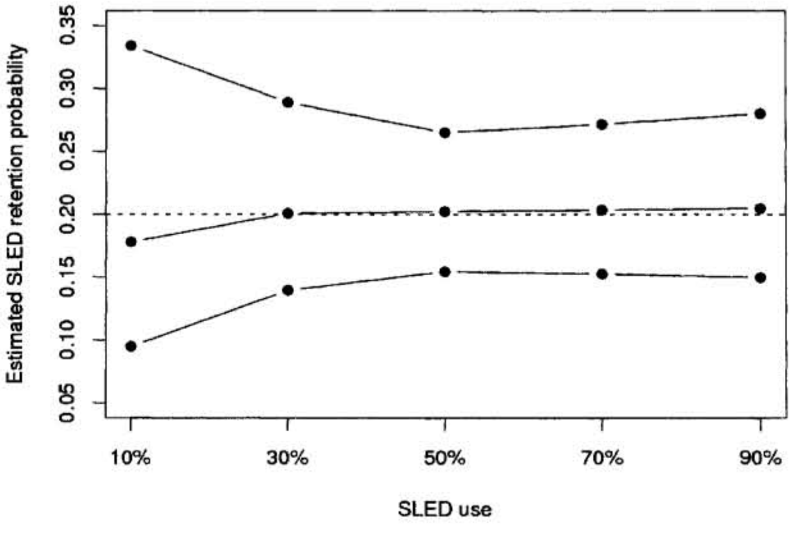

Figure 16: The model estimated SLED retention probability (mean and 95% c.i.), as a function of the assumed level of SLED coverage.