Robustness of the penguin population model estimator

William Robinson and Doug Butterworth

MARAM (Marine Resource Assessment and Management Group)
Department of Mathematics and Applied Mathematics,
University of Cape Town

November 2010

Introduction

The robustness of an estimator may be investigated by assuming that the "true" operating model is in fact different (for example, a constant may have a different value or a key relationship may have a different functional form to that assumed by the estimator). Data generated from this "reality" are then fitted using the estimator. The distributions of the results are then compared to the corresponding "true" values of the operating model.

Method and results

As an initial sensitivity analysis, two variations of the logistic transformation operating model described in MARAM IWS/DEC10/PA/P6 are tested:

1. the value of the relative juvenile detectability p_{J} is changed from 1.0 to 0.9, and
2. the dependency on sardine biomass is changed from proportional to $\ln B$ to proportional to $\sqrt{\ln B}$.

In operating model 2, the biomass index was changed to

$$B_y = I_y / \min\{I\} \tag{1}$$

so that $\ln B \geq 0$.

In each case, the maximum penalized likelihood estimates of the adult female moult counts $\hat{N}_y$ and the proportion of immature birds in the moult counts $\hat{J}_y$ under these alternative assumptions are input to the data-generation routine. Pseudo-datasets are then generated from the best estimates as follows:

$$\begin{aligned} N_{y,i}^{\text{pseudo}} &= \hat{N}_y e^{X_{y,i}} & X_{y,i} &\sim N\left(0, \sigma_M^2\right) \\ J_{y,i}^{\text{pseudo}} &= \hat{J}_y e^{Y_{y,i}} & Y_{y,i} &\sim N\left(0, \sigma_J^2\right) \end{aligned} \tag{2}$$

The standard deviations are the same as those used in the estimator: $\sigma_M = 0.2$ and $\sigma_J = 0.1$. Note that the lower estimates from the model fit residuals are not used because they are negatively biased as a result of the relatively few degrees of freedom for the penalized maximum likelihood estimate used.

The estimator then finds the best fit to each pseudo-dataset. The results for the lower value of p_J are shown in Figure 2 to Figure 4, and the results for the different functional form for the prey abundance dependency are shown in Figure 8 to Figure 10. The distributions of the estimated μ and η parameters, and the spread of estimated values for the time series of adult female moult counts, juvenile proportions in the counts, annual survival and reproductive success are displayed.

This process is identical to a standard parametric bootstrap procedure, which is implemented here to check the robustness of the estimator.

Discussion

Of key interest are the μ parameters which define the relationships in the model between demographic parameters and prey abundance. Figure 1 and Figure 7 show the fits of the biomass indices to the annual survival rate. Figure 2 and Figure 8 show that the bootstrap distributions quite different to the "true" values in most cases. For the lower value of p_J, the "true" annual survival rate is lower than that which the estimator recovers (Figure 3), while the reproductive success rate is higher (Figure 4). Of more relevance to population projections is the series of moult counts which the estimator recovers. For both operating models, there is little difference between the "true" counts and the median estimated values (Figure 6 and Figure 12).

These biases in the μ parameter estimates are not unexpected. The key question is how they affect the quantities of ultimate interest: the impact of different pelagic fish harvesting strategies. That will be addressed in a further submission.

Note that as in MARAM IWS/DEC10/PA/P7 this process has considered the effects of observation errors in the moult count data only. A more comprehensive test would need to consider also alternative realisations of the random effects for the underlying operating model.

Reference

MARAM IWS/DEC10/PA/P6. Robinson W, Butterworth DS. 2010. Penguin population models for Robben Island.

MARAM IWS/DEC10/PA/P7. Robinson W, Butterworth DS. 2010. Checking the penguin population model for bias.

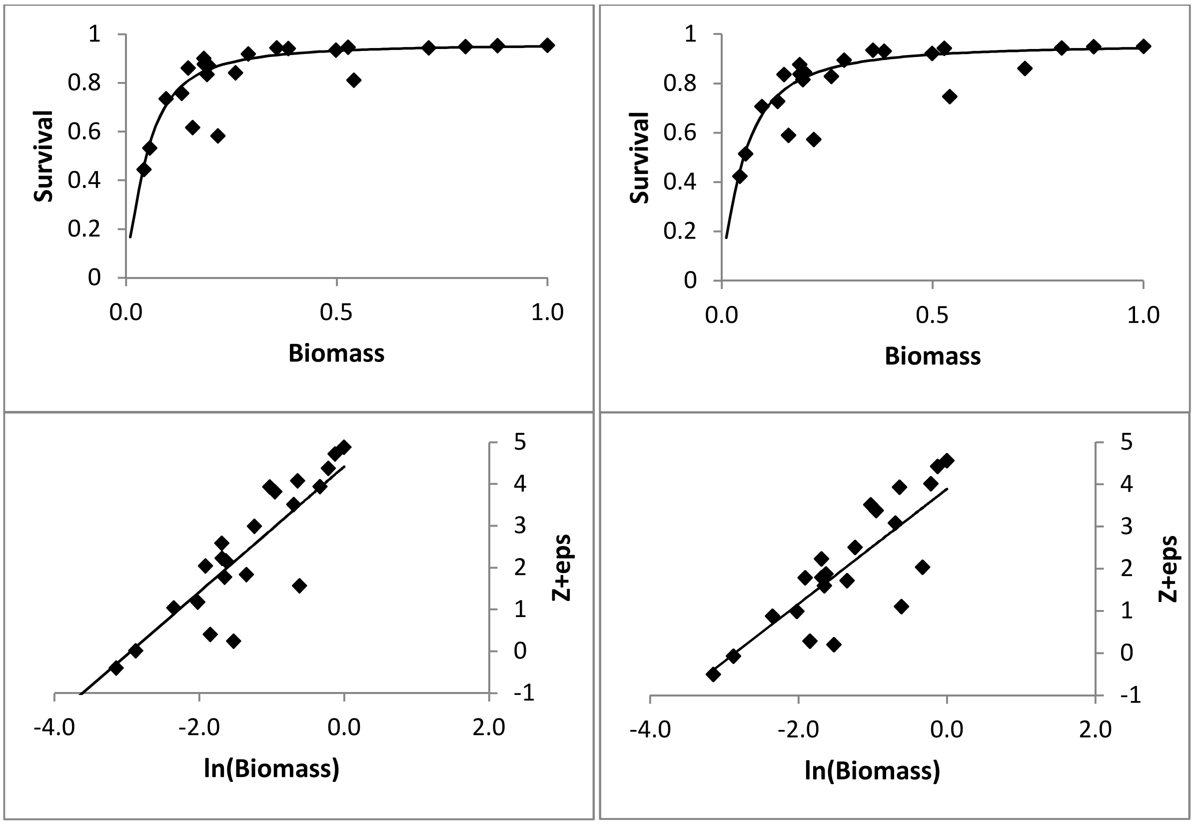

Figure 1: Fits of annual survival to sardine biomass for: (left) the base case $\left(p_J = 1.0\right)$ and (right) the alternative $\left(p_J = 0.9\right)$.

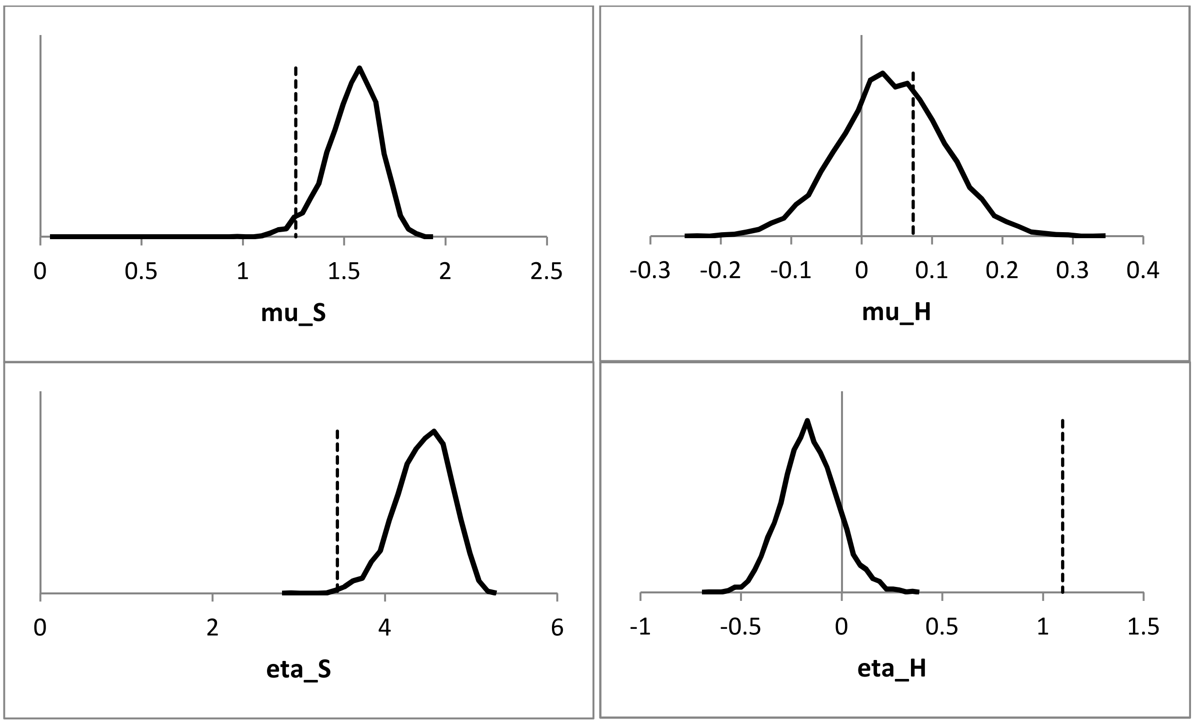

Figure 2: Distributions of the parameter estimates for the relationships with fish abundance obtained from the bootstrapped data. The "true" values for the first alternative operating model $\left(p_J = 0.9\right)$ are indicated by dashed lines.

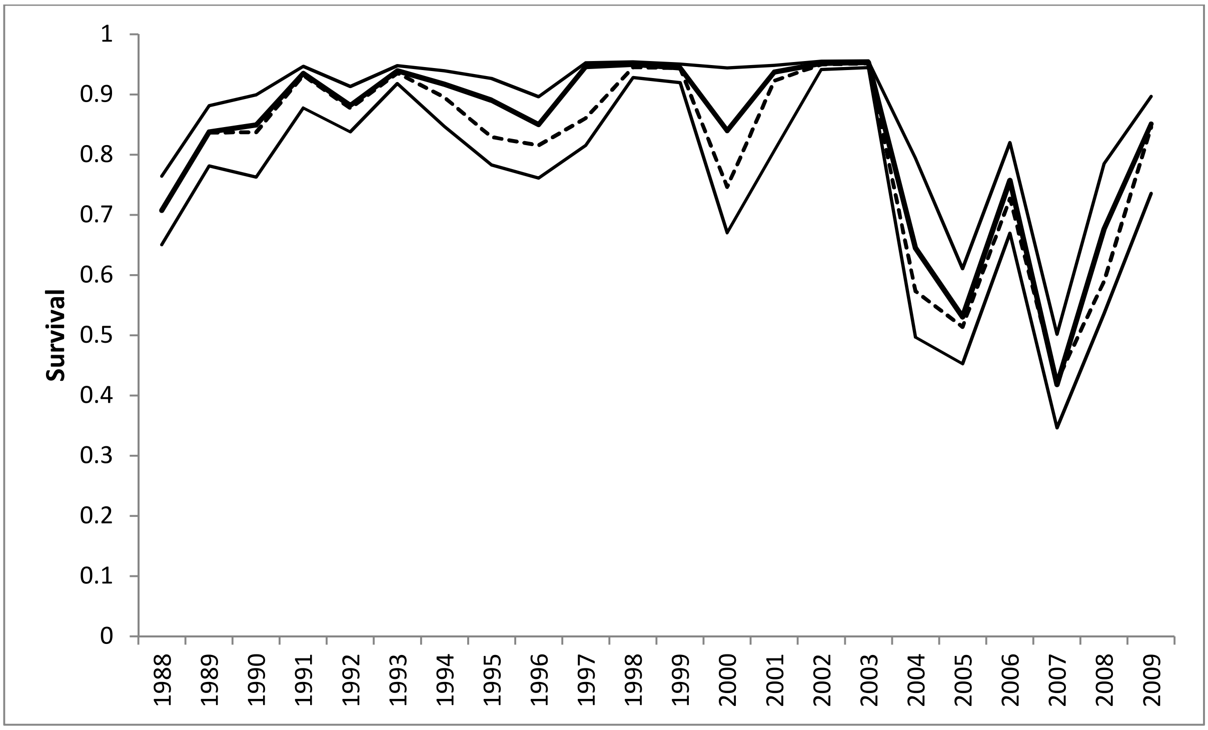

Figure 3: 5th percentile, median and 95th percentile of the estimated annual adult penguin survival rates from the bootstrapped data. The dashed line indicates the maximum penalized likelihood estimates for the operating model with $p_J = 0.9$.

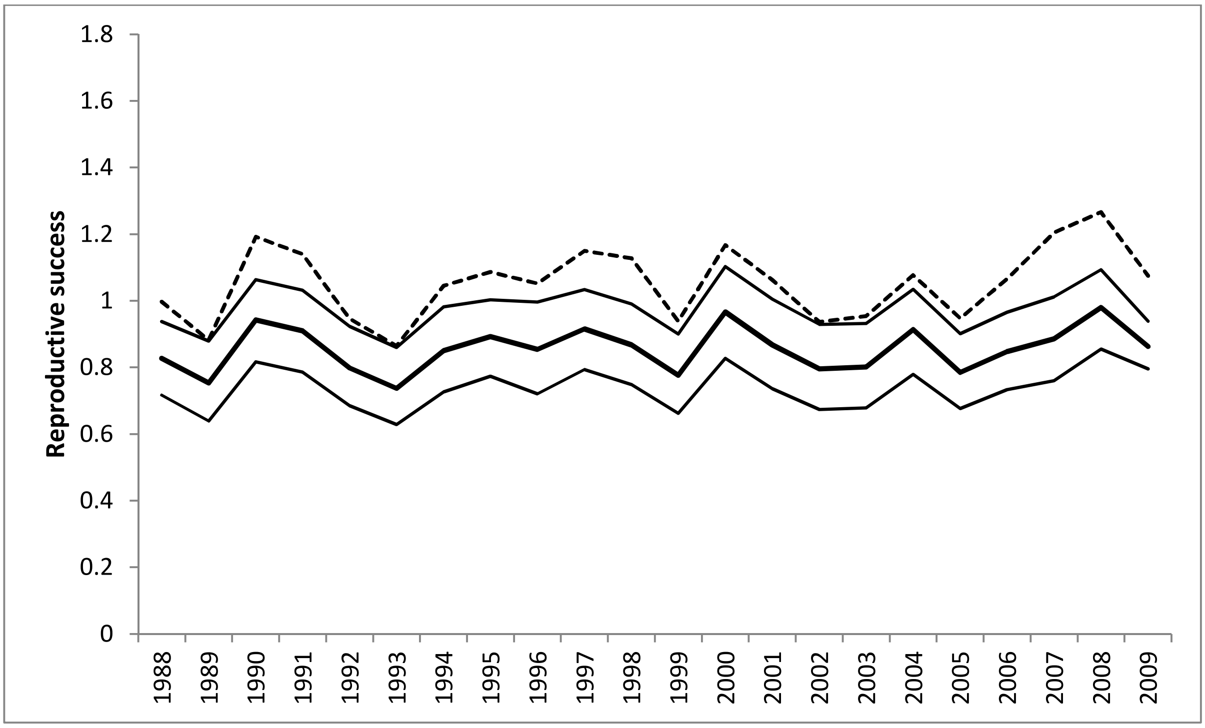

Figure 4: 5th percentile, median and 95th percentile of the estimated annual penguin reproductive success rates from the bootstrapped data. The dashed line indicates the maximum penalized likelihood estimates for the operating model with $p_J = 0.9$.

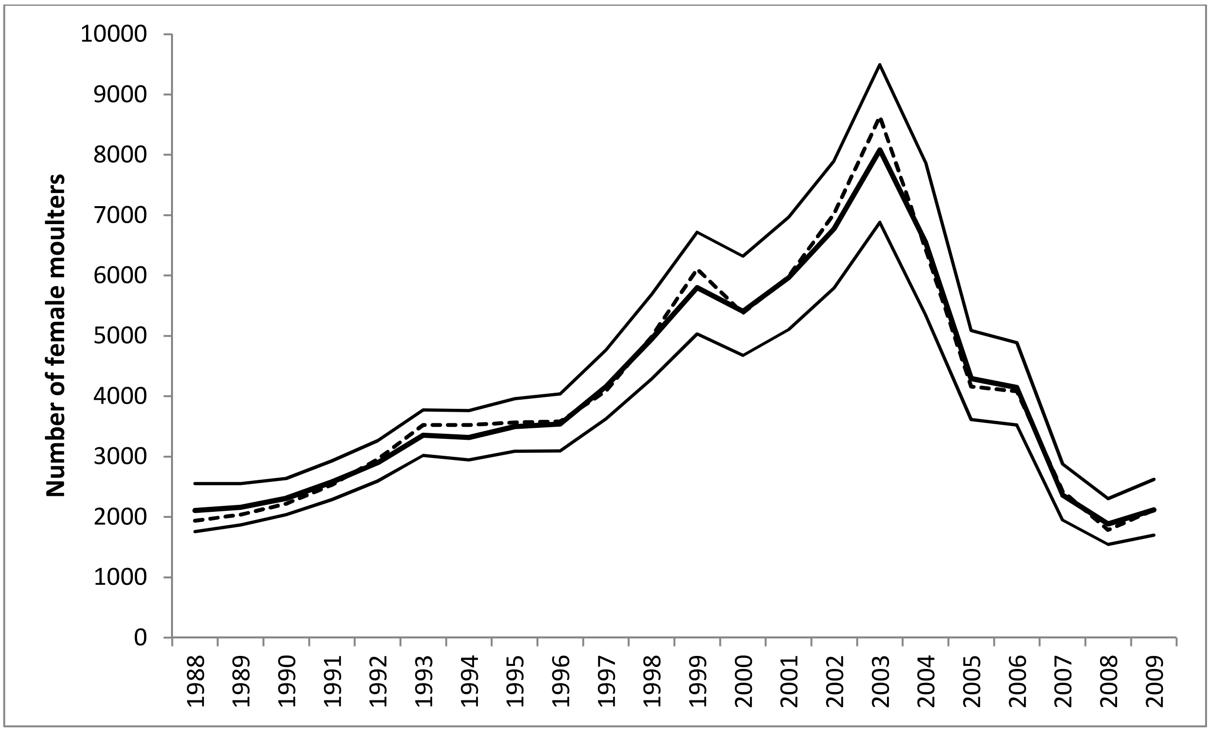

Figure 5: 5th percentile, median and 95th percentile of the estimated adult female moult counts from the bootstrapped data. The dashed line indicates the maximum penalized likelihood estimates for the operating model with $p_J = 0.9$.

Figure 6: 5th percentile, median and 95th percentile of the estimated proportion of immature penguins in the moult counts from the bootstrapped data. The dashed line indicates the maximum penalized likelihood estimates for the operating model with $p_J = 0.9$.

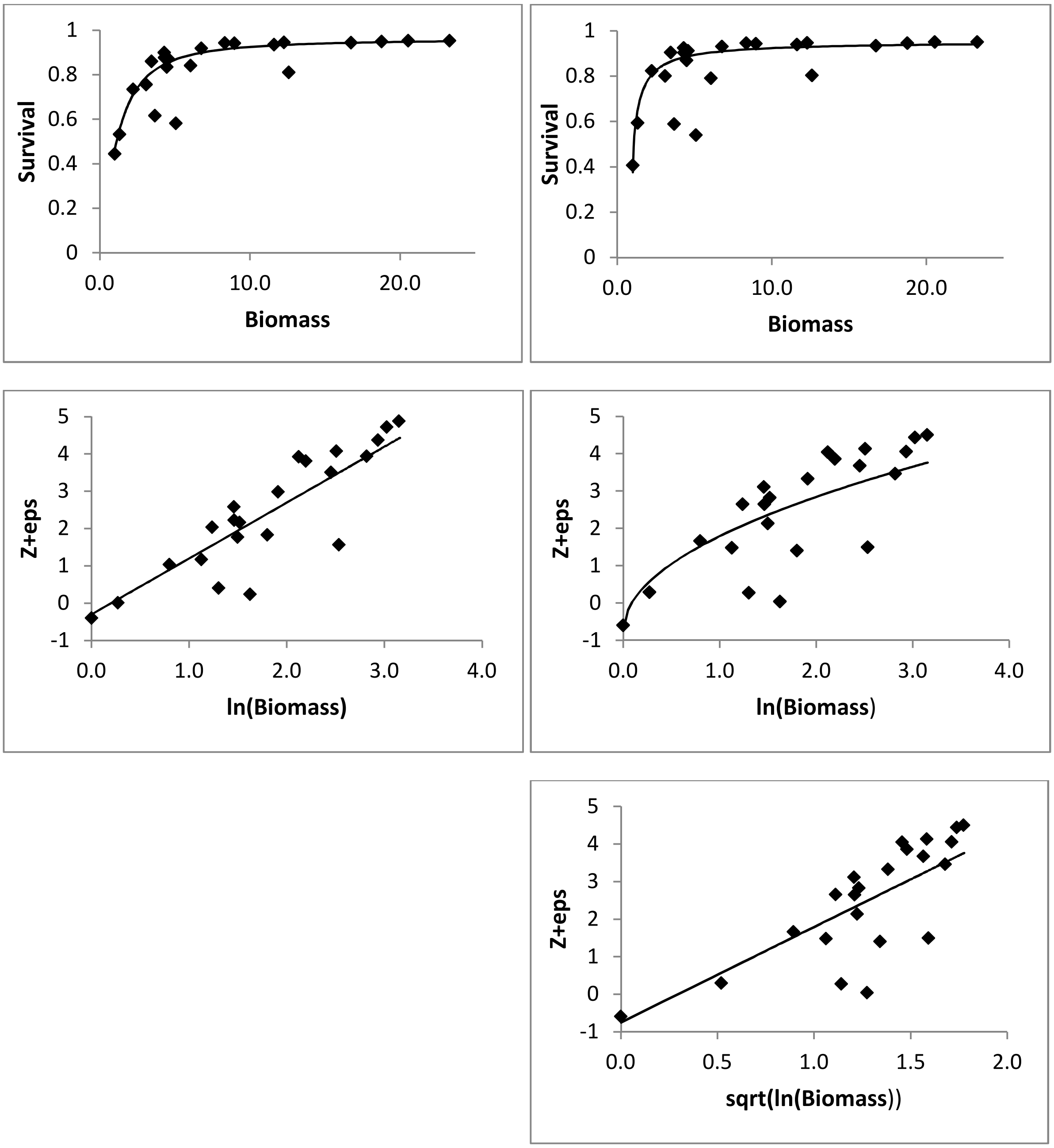

Figure 7: Fits of annual survival to sardine biomass for: (left) the base case (Z proportional to the logarithm of the biomass index) and (right) the alternative (Z proportional to the square root of the logarithm of the biomass index).

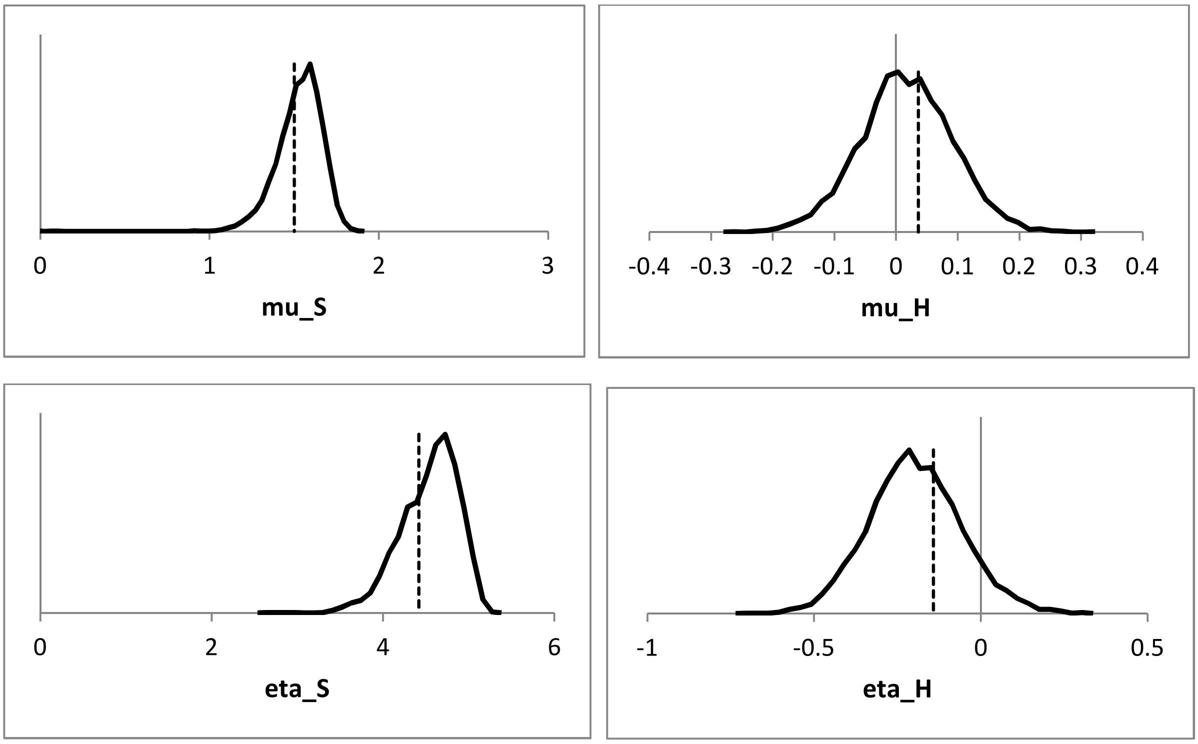

Figure 8: Distributions of the parameter estimates for the relationships with fish abundance obtained from the bootstrapped data. The maximum penalized likelihood estimates for the operating model with an alternative biomass dependence are indicated by dashed lines.

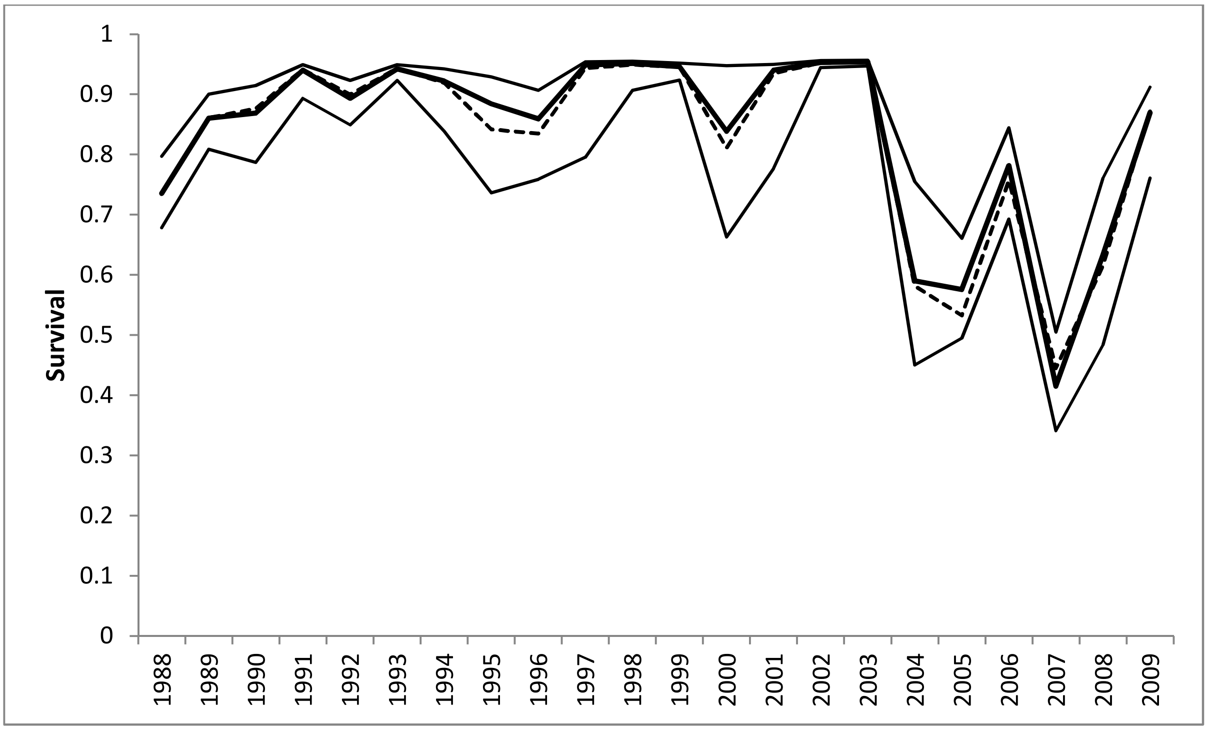

Figure 9: 5th percentile, median and 95th percentile of the estimated annual adult penguin survival rates from the bootstrapped data. The dashed line indicates the maximum penalized likelihood estimates for the operating model with an alternative biomass dependence.

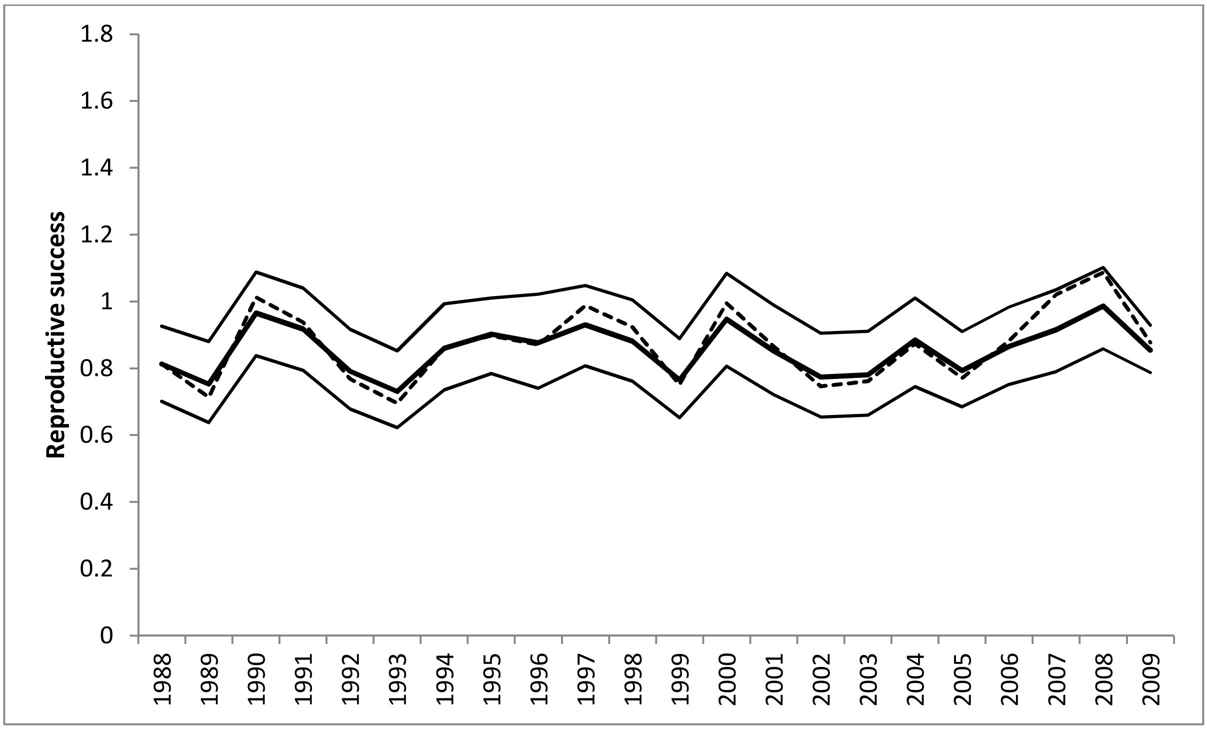

Figure 10: 5th percentile, median and 95th percentile of the estimated annual penguin reproductive success rates from the bootstrapped data. The dashed line indicates the maximum penalized likelihood estimates for the operating model with an alternative biomass dependence.

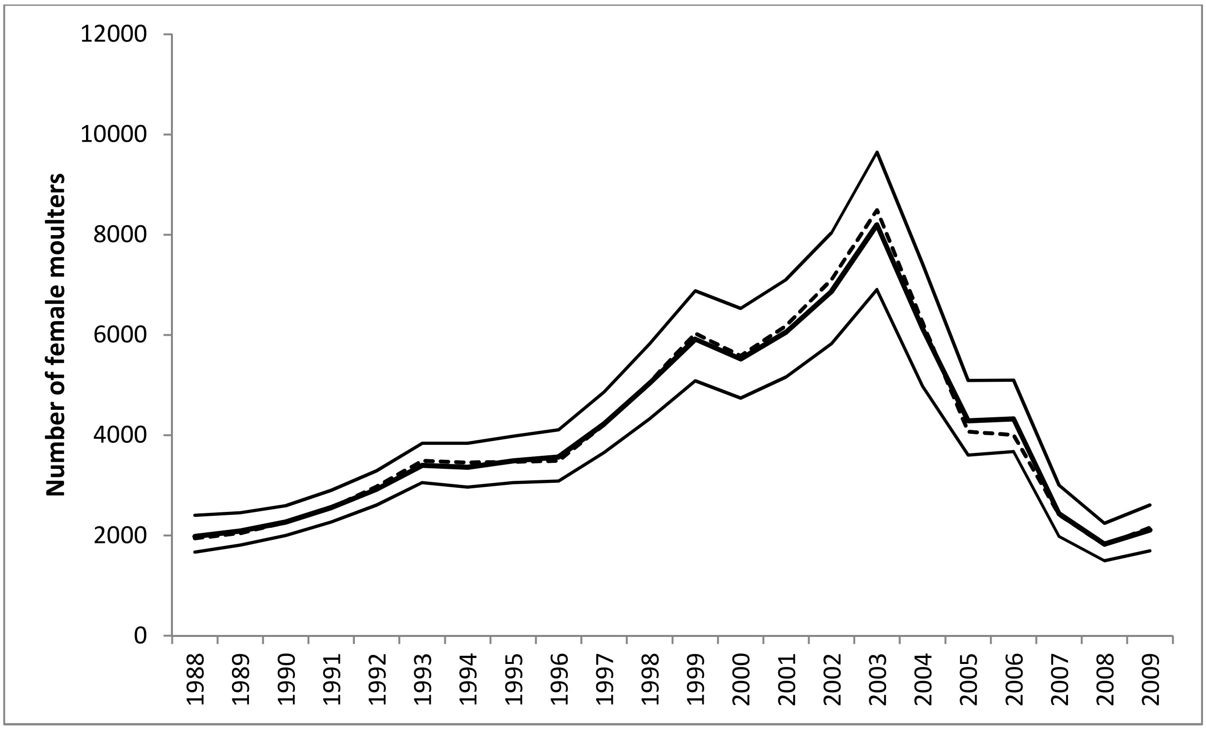

Figure 11: 5th percentile, median and 95th percentile of the estimated adult female moult counts from the bootstrapped data. The dashed line indicates the maximum penalized likelihood estimates for the operating model with an alternative biomass dependence.

Figure 12: 5th percentile, median and 95th percentile of the estimated proportion of immature penguins in the moult counts from the bootstrapped data. The dashed line indicates the maximum penalized likelihood estimates for the operating model with an alternative biomass dependence.