Full description of the Robben Island Penguin–Fish interaction model

William Robinson[1] and Doug Butterworth

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

November 2011

Introduction

Work on the penguin–fish interaction model has progressed to the point where a reasonable new base case for Robben Island has been achieved.

Data and assumed parameters

Table 1 lists the moult count data to which the model is fitted. An overview of the data collection process and the method for deriving the moult count time series has been reported (Robinson & Butterworth, 2010). The time series has been updated with the 2011 counts.

Estimates of the direct impact of the two major oil spills are given in Table 2, based on published information (Crawford et al., 2000; Underhill et al., 1999).

Table 4 provides a composite list of model parameter values assumed on input.

Estimates for the abundances of sardine and anchovy come from the acoustic surveys which take place in May and November each year (de Moor, Butterworth, & Coetzee, 2008).

Notes on the model

A fundamental component of the model is the assumed relationship between pelagic fish abundance and penguin adult mortality. The process of choosing the statistical distribution assumed for the variability of penguin mortality about the value suggested by the relationship with prey abundance has been through several iterations. In particular, the beta distribution option had to be discarded, since the unimodal beta distribution can only have a very small variance when the mean is close to the boundary. This gives a high weight to years in which survival values want to go close to the maximum, which itself drives estimates towards the upper boundary. This is undesirable since it implies that when prey biomass is high, survival can only be very close to the maximum. An alternative approach, described below, has been explored and found to work satisfactorily.

Sightings of tagged birds over the period 1989–2010 have now been included using a multinomial likelihood. Over-dispersion has been estimated using the program MARK applied to the tag-recapture data in isolation.

---

[1]william.robinson@uct.ac.za

Immigration of three year old birds to Robben Island has now been allowed over the period 1989–1999. This can now be estimated given tag-recapture data over this period (not previously available) which remove the confounding between immigration and survival rates that otherwise applies. It is plausible that immigration, likely from Dyer Island and Dassen Island, occurred over this period.

Bayesian integration was executed using the MCMC algorithm implemented in AD Model Builder. A chain of length 5,500,000 was run, discarding the first 500,000 and saving every 1000th iteration. Thus a sample of 10,000 was used to estimate posterior distributions. Initial checks revealed nothing to suggest any problems with convergence.

Model description

Basic dynamics

The model considers the number of female penguins $N_{y,a}$ at the start (1 January) of year y of age a at Robben Island. The initial population size (at the start of year 1988) N_0 and its age structure are:

$$N_{y_0,a} = \begin{cases} N_0 e^{-a\lambda} & \text{for } 1 \le a < A \\ \dfrac{N_0 e^{-a\lambda}}{1-e^{-\lambda}} & \text{for } a = A \end{cases} \qquad (1)$$

where A is the plus-group age. Both N_0 and λ are parameters whose values are estimated.

The following equations describe the population trajectory:

$$N_{y+1,a} = \begin{cases} \dfrac{1}{2} H_y \left(\sum_{a=a^*}^{A} N_{y,a} \right) S_y^{4/12} \left(1 - p_y^{\text{oil}}\right) & \text{if } a = 1 \\ N_{y,a-1}^* S_y & \text{if } a = 2 \\ N_{y,a-1}^* S_y + I_y & \text{if } a = 3 \\ N_{y,a-1}^* S_y & \text{if } a = 4 \\ \left(N_{y,a-1}^* + N_{y,a}^* \right) S_y & \text{if } a = A \end{cases} \qquad (2)$$

with

$$N_{y,a}^* = N_{y,a} \left(1 - \frac{N_y^{\text{oil}}}{\left(\sum_{a=1}^{A} N_{y,a} \right) S_y^{m_y/12}} \right) \qquad (3)$$

where

S_y is the adult (post 1 January of first year of life) annual survival rate in year y (taken to the power 4/12 in the equation for $a = 1$ since the peak of the breeding season is assumed to be on 1 May),

H_y is the annual reproductive success (number of chicks per mature female reaching 1 January of the year following birth, where 50% of these chicks are assumed to be female),

a^* is the age at which the penguins first attempt to breed (taken here to be $a^* = 4$),

p_y^{oil} is the proportion of chicks estimated to have died as a result of an oil spill in year y,

I_y is the number of penguins (all assumed to be of age 3) immigrating to Robben Island in year y (earlier investigations allowing more flexible variation over time suggested that it is appropriate to estimate this as three constant levels for the periods 1989–1991, 1992–1994 and 1995–1999),

N_y^{oil} is the number of juvenile and adult penguins estimated to have died as a result of oiling in year y, and

m_y is the month in which the oil spill occurred in year y.

See Table 2 for the values used for p_y^{oil}, N_y^{oil} and m_y.

Population model

Both the annual adult survival rate S_y and the annual reproductive success H_y are assumed to depend on some function of prey biomass (the deterministic effect), but to be influenced also by some noise (random effects). For reproductive success, the estimates are drawn from a beta distribution which ensures that biologically plausible bounds are respected. For adult survival, the alternative approach below was developed and has been found to work satisfactorily.

Adult survival depends on the normalized annual biomass levels $B_{S,y}$, where the time series I_S is some function of the sardine and anchovy November spawner biomass and May recruit biomass survey results for particular spatial regions:

$$B_{S,y} = I_{S,y} \big/ \max\{I_S\} \tag{4}$$

Similarly, reproductive success depends on a similar index of biomass level:

$$B_{H,y} = I_{H,y} \big/ \operatorname{mean}\{I_H\} \tag{5}$$

In the base case implemented here, B_S is taken to be the sardine November survey spawner biomass west of Cape Agulhas (strata A–C). These values are plotted in Figure 1, with the corresponding biomass in tons on the right-hand axis. B_H is anchovy recruit survey biomass west of Cape Infanta.

Adult survival

The rates of annual adult survival S_y and annual natural mortality M_y are related as

$$S_y = e^{-M_y} \tag{6}$$

where M_y is modelled as follows:

$$M_y = M_{\min} + f_S\left(B_{S,y}\right)e^{X_y} \tag{7}$$

and X_y is distributed $N\left(0,\sigma_y^2\right)$ with

$$\sigma_y = \sqrt{e^{\left[\tilde{\sigma}/f_S\left(B_{S,y}\right)\right]^2} - 1} \tag{8}$$

Thus there is a log-normal random effect, but since the σ_y depend on the biomass $B_{S,y}$, the M_y distributions will all have exactly the same standard deviation. This is appropriate since then the data related to each year receive roughly equal weighting, and, when projecting, high resource biomass does not force low penguin mortality.

The prior added to the negative log likelihood for each year to reflect assumptions made above for the X_y parameters is:

$$P_S = \sum_y \left[\ln \sigma_y + \frac{1}{2}\left(\frac{X_y}{\sigma_y}\right)^2 \right] \tag{9}$$

An additional penalty term ("prior") ensures that the annual mortality rates are evenly distributed about the curve relating mortality and biomass, specifically the sum of the residuals is forced to zero (this was found to aid estimation stability):

$$P_{\text{B-M}} = 10^5 \left\{ \sum_y \left[M_y - \left(M_{\min} + f_S\left(B_{S,y}\right) \right) \right] \right\}^2 \tag{10}$$

Biomass–mortality relationship 1

The first candidate biomass–mortality relationship evaluated was defined as:

$$f_S\left(B_{S,y}\right) = M_{\min}^* + \left(a + \frac{B_{S,y}}{b} \right)^{-n} \tag{11}$$

The motivation for including the $M_{\min}^*$ term is that, when calculating M_y in equation (7), the possibility is excluded that the term additional to $M_{\min}$ could go to zero when biomass is high. This would be problematic as then σ_y would go infinite in equation (8), rendering the associated data point redundant because of its infinite variance. The lower bound on achieved M_y remains $M_{\min}$,

but the lower bound on the median of its distribution is $M_{\min} + M_{\min}^{*}$. The base case model uses the following constant values: $\tilde{\sigma} = 0.1$, $M_{\min} = 0.04$ and $M_{\min}^{*} = 0.02$.

Instead of a and b, equation (11) can be re-parameterised in terms of the values of $f_S\left(B_{S,y}\right)$ at $B_S = 0.1$ and $B_S = 1$ in order to introduce estimable parameters which are more orthogonal (i.e. the new parameters have low correlation) which assists with statistically stable estimation. If $U = f_S\left(B_S = 1\right)$ and $V = f_S\left(B_S = 0.1\right) - f_S\left(B_S = 1\right)$, then

$$b = \frac{0.9}{\left(U - M_{\min}^{*}\right)^{-1/n} - \left(U + V - M_{\min}^{*}\right)^{-1/n}} \tag{12}$$

$$a = \left(U - M_{\min}^{*}\right)^{-1/n} - \frac{1}{b} \tag{13}$$

The estimable parameters are then U, V and n, where sensibly $V \geq 0$.

A concern arising from analyses with this functional form is that it turned out that the likelihood could always be increased by raising the value of the power parameter n, leading to difficulty in choosing an appropriate prior for n. This indicates that the functional form is perhaps not the most appropriate. Hence an alternative form, described in the following section, has been adopted.

Biomass–mortality relationship 2

The power relationship described above has been discarded in favour of a series of connected straight lines. The relationship is assumed to be constant above the biomass level $B_S = 0.4$. Below this level of prey abundance, mortality increases as biomass decreases. (It was found that allowing more complex behaviour for $B_S > 0.4$ did not improve the model fit significantly.) The function is set up to ensure that the gradient increases as biomass decreases, as follows:

$$\begin{aligned}
f_S(1.0) &= M_1 \\
f_S(0.4) &= M_1 \\
f_S(0.3) &= f_S(0.4) + M_2 \\
f_S(0.2) &= 2f_S(0.3) - f_S(0.4) + M_3 \\
f_S(0.1) &= 2f_S(0.2) - f_S(0.3) \\
f_S(0.0) &= 2f_S(0.1) - f_S(0.2)
\end{aligned} \tag{14}$$

The estimable parameters are M_1, M_2 and M_3, each of which must be positive. In theory, additional M parameters could be added to the final two lines above, but it was found that these did not improve the fit significantly. Linear interpolation is used to calculate the mortality at intermediate biomass values.

Both form 1 and form 2 of the function $f_S(B_S)$ as estimated in fitting the model are shown in Figure 2.

Reproductive success

The parameters $\bar{H}_y^* = f_H\left(B_{H,y}\right)$ are estimated on the interval $[0,1]$. These are the transformed predicted annual reproductive success rates $\bar{H}_y$ which fall in the interval $[0, H_{max}]$:

$$\bar{H}_y = H_{max}\bar{H}_y^* \tag{15}$$

The assumed relationship between $\bar{H}_y^*$ and the fish abundance index $B_{H,y}$ is currently assumed to be a constant given that earlier analyses have not provided any indication of dependence:

$$f_H\left(B_{H,y}\right) = h \tag{16}$$

but alternative functional forms may be considered in future.

The parameters H_y^* are estimated for each year on the interval $[0,1]$ and then transformed to the range $[0, H_{max}]$ to obtain the annual reproductive success rates H_y as follows:

$$H_y = H_{max}H_y^* \tag{17}$$

We assume that the H_y^* parameters are beta-distributed about the predicted values $\bar{H}_y^*$. Setting $\kappa_{H,y} = \bar{H}_y^*\left(1-\bar{H}_y^*\right)\big/\sigma_H^2 - 1$, the parameters of this beta distribution $\alpha_{H,y}$ and $\beta_{H,y}$ are:

$$\alpha_{H,y} = \bar{H}_y^*\kappa_{H,y} \qquad \beta_{H,y} = \kappa_{H,y}\left(1-\bar{H}_y^*\right) \tag{18}$$

The prior added to the negative log likelihood for each year, which reflects the assumption that H_y^* is beta-distributed, is:

$$P_{H,y} = -\ln\left(\frac{\Gamma\left(\alpha_{H,y}+\beta_{H,y}\right)}{\Gamma\left(\alpha_{H,y}\right)\Gamma\left(\beta_{H,y}\right)}\right) - \left(\alpha_{H,y}-1\right)\ln H_y^* - \left(\beta_{H,y}-1\right)\ln\left(1-H_y^*\right) \tag{19}$$

An additional penalty term ensures that the annual reproduction success rates are evenly distributed about the assumed relationship curve, i.e. the sum of the residuals is forced to zero (this was found to aid estimation stability):

$$P_{B\text{-}R} = 10^5\left\{\sum_y\left[H_y - \bar{H}_y\right]\right\}^2 \tag{20}$$

Fitting to moult counts

The population model is fitted to annual moult count data for both adult and juvenile birds (Table 1) by taking them into account through the following negative log-likelihood functions:

$$-\ln L_{\mathrm{M}} = \sum_y \left[\ln \sigma_{\mathrm{M},y} + \frac{1}{2\sigma_{\mathrm{M},y}^2} \left(\ln N_y^{\mathrm{obs}} - \ln N_y^{\mathrm{model}} \right)^2 \right]$$

$$-\ln L_{\mathrm{J}} = \sum_y \left[\ln \sigma_{\mathrm{J}} + \frac{1}{2\sigma_{\mathrm{J}}^2} \left(\ln J_y^{\mathrm{obs}} - \ln J_y^{\mathrm{model}} \right)^2 \right]$$

(21)

$\sigma_{\mathrm{M},y}$ and σ_{J} are respectively the standard errors of the logarithms of the adult moult counts and juvenile proportions of these counts about their true values (i.e. these reflect observation errors). The terms in the likelihood for the adult moult counts are weighted according to the CVs of the observations (see Table 1) with

$$\sigma_{\mathrm{M},y} = \sqrt{\mathrm{CV}\left(N_y^{\mathrm{obs}} \right)^2 + 0.05^2 + \sigma_{\mathrm{add}}^2}$$

(22)

where the additional variance σ_{add}^2 is an estimable parameter, with the true additional variance forced to be at least 0.05^2. The value of σ_{J} is fixed at $\sigma_{\mathrm{J}} = 0.1$.

$N_y^{\mathrm{model}} = q_{\mathrm{M}} \sum_{a=2}^{A} N_{y,a} S_y^{11/12}$ is the number of female birds in adult plumage (aged 2 and over) counted in year y, where the peak of the counting season is the end of November, and q_{M} is the proportion of these birds susceptible to observation (assumed here to be $q_{\mathrm{M}} = 0.9$).

$J_y = \dfrac{p_{\mathrm{J}} N_{y,1}}{\sum_{a=1}^{A} N_{y,a}}$ is the proportion of juvenile birds in the model at the time of the moult count (note that $S_y^{11/12}$ factors in numerator and denominator cancel), with p_{J} being the detectability of juvenile moulters relative to adults in the counting process, which is assumed here to be $p_{\mathrm{J}} = 1$.

N_y^{obs} is the number of female adult moulters observed in year y (taken to be one half of the total adult moulters counted).

J_y^{obs} is the observed proportion of moulters in immature plumage counted in year y.

Fitting to tag data

In addition to the moult counts, tag data for the period 1989–2010 have been incorporated into the analysis. This provides an independent estimate of annual survival rates and allows for the estimation of immigration. The multinomial likelihood of the encounter histories is calculated through the estimation of re-sighting probabilities and survival rates for each year. This is the same method as used in program MARK. A variation on the standard MARK procedure which has been included in the analysis is the estimation of a "transient" factor relating to birds tagged which are never seen again, modelled as additional mortality in the year following tagging to reflect emigration. This factor is assumed to be equal for all years, except for the years of the major oil spills

(*Apollo Sea* in 1994 and *Treasure* in 2000) for which separate values are estimated, as these birds are more likely to be linked to Robben Island than would be the case for normal "transients".

The log-likelihood function for the multinomial distribution, ignoring the constant multinomial coefficient, is:

$$\ln L_{\text{T-R}} = \frac{1}{\hat{c}} \sum_{j=1}^{m} n_j \ln \hat{p}_j \tag{23}$$

where $\hat{p}_j$ is the estimated probability of the j th encounter history occurring and n_j is the number of times which that history has been observed. The number of unique encounter histories is m. In order to account for over-dispersion, the likelihood is scaled by a factor of $1/\hat{c}$, whose value was obtained from analysing the data using MARK.

Objective function

The overall log posterior is thus:

$$-\ln P_{\text{post}} = -\ln L_{\text{M}} - \ln L_{\text{J}} - \ln L_{\text{T-R}} + P_S + P_{\text{B-M}} + P_H + P_{\text{B-R}} \tag{24}$$

Note that all other priors are uniform so do not contribute (Table 5).

Sensitivities

Most of the base case model sensitivities which were tested consider variation in values for parameters which were fixed on input to the model. In these cases the base case value is given in parenthesis below. The following variations were tested:

1. Mortality–sardine biomass relationship 1.
2. Expected values for numbers of tag re-sightings each year forced to match observations.
3. Relative detectability of juveniles $p_J = 0.9 \ (p_J = 1.0)$.
4. Age of first breeding attempt $a^* = 3 \ (a^* = 4)$.
5. Age of first breeding attempt $a^* = 5 \ (a^* = 4)$.
6. Standard error of the logarithms of the juvenile proportions $\sigma_J = 0.2 \ (\sigma_J = 0.1)$.
7. Variability about the biomass–mortality relationship $\tilde{\sigma} = 0.05 \ (\tilde{\sigma} = 0.1)$.
8. Variability about the biomass–mortality relationship $\tilde{\sigma} = 0.2 \ (\tilde{\sigma} = 0.1)$.
9. Maximum breeding success rate $H_{\max} = 1.5 \ (H_{\max} = 1.8)$.
10. Proportion of moulters susceptible to observation $q = 0.8 \ (q = 0.9)$.
11. Proportion of moulters susceptible to observation $q = 1.0 \ (q = 0.9)$.

The reason for the inclusion of sensitivity 2 above is that, as evident from Figure 3, the new base case has not changed earlier results of greater numbers of re-sightings than predicted by the model over most years post-2000.

Results

Table 5 provides a composite list of model parameters which are estimated and the priors assumed for them. Parameter estimates at the joint posterior mode and the medians and 90% probability intervals of the Bayesian posterior distributions are given for the new base case.

Table 6 lists the results of the sensitivity tests. As the primary purpose of this modelling exercise is to predict future trends in penguin abundance in relation to future sardine biomass levels, these results have been expressed in terms of the penguin trends for the next 10 years and how they relate to those for the new base case for levels of future sardine abundance close to those required to sustain penguin numbers.

The tables are followed by various illustrative plots.

Discussion

Generally the fits to the data and the residual patterns shown in Figure 3 to Figure 5 appear acceptable. A particular feature warranting further attention is the tendency for observed re-sightings to exceed expected numbers after the year 2000. This has been investigated in sensitivity test 2.

Note estimates of annual immigration over the 1989–1994 period with 90% probability intervals ranging from about 100 to 1200 (see Table 5). This range seems reasonably compatible with numbers of penguins counted at Dyer and Dassen Islands (the likely origins of these immigrants) over the period in question, though perhaps an upper bound should be introduced.

When expected numbers of banded penguin re-sightings are forced to equal the observed values, the overall fit deteriorates (see Figure 12 (b)). Figure 13 compares estimated annual adult survival for the new base case and the variation forcing the tag data fit. Note that in the variation, the estimates of survival are higher over the period 2004–2008 which is in conflict with the information provided by the moult counts.

Figure 14 shows that the fits for the two mortality–biomass functional forms tested are almost identical, while there are slight differences in the projections. The new functional form predicts more of a decline in penguin numbers at $B_S = 0.2$, but much the same increase in numbers at $B_S = 0.3$.

The input parameter which has the greatest sensitivity at low sardine biomass levels is $\tilde{\sigma}$ which is a measure of the variability about the biomass–penguin mortality relationship. This is not an unexpected result, and the value adopted for $\tilde{\sigma}$ warrants further consideration. Sensitivity to the other input parameter value changes is slight, except perhaps for the age at which breeding commences.

Future work

Now that a satisfactory new base case model has been attained, the following steps will be pursued:

1) linking the model with the updated OMP operating models to explore the consequences for future penguin abundance of alternative pelagic fish harvesting levels; and

2) extending the model first to some other Western Cape colonies in isolation, and then in combination.

References

Crawford, R. J. M., Davis, S. A., Harding, R. T., Jackson, L. F., Leshoro, T. M., Meÿer, M. A., Randall, R. M., et al. (2000). Initial impact of the Treasure oil spill on seabirds off western South Africa. *South African Journal of Marine Science*, *22*, 157-176.

de Moor, C. L., Butterworth, D. S., & Coetzee, J. C. (2008). Revised estimates of abundance of South African sardine and anchovy from acoustic surveys adjusting for echosounder saturation in earlier surveys and attenuation effects for sardine. *African Journal of Marine Science*, *30*(2), 219-232. doi:10.2989/AJMS.2008.30.2.2.552

Robinson, W. M. L., & Butterworth, D. S. (2010). *A proposed new method for obtaining penguin moult count estimates* (pp. 1-13).

Underhill, L. G., Bartlett, P. A., Baumann, L., Crawford, R. J. M., Dyer, B. M., Gildenhuyst, A., Nel, D. C., et al. (1999). Mortality and survival of African Penguins Spheniscus demersus involved in the Apollo Sea oil spill: an evaluation of rehabilitation efforts. *Ibis*, *141*, 29-37.

Tables

Table 1: Counts of birds moulting at Robben Island with CVs and the proportions of counts which are juveniles.

Year	Number of adult female moulters	CV	Proportion of juveniles
1988/1989	1733	7.9	0.188
1989/1990	1698	7.4	0.200
1990/1991	2362	6.1	0.162
1991/1992	2470	6.0	0.245
1992/1993	3272	5.3	0.199
1993/1994	3958	4.8	0.166
1994/1995	3966	4.8	0.147
1995/1996	3337	5.4	0.174
1996/1997	3650	8.0	0.239
1997/1998	4278	4.4	0.218
1998/1999	4562	4.3	0.231
1999/2000	5847	4.0	0.194
2000/2001	6442	3.8	0.158
2001/2002	6181	8.6	0.180
2002/2003	7261	3.9	0.191
2003/2004	8488	3.6	0.164
2004/2005	6221	4.0	0.168
2005/2006	3830	4.9	0.254
2006/2007	3226	6.0	0.226
2007/2008	2579	6.4	0.213
2008/2009	1872	7.2	0.325
2009/2010	2105	19.1	0.291
2010/2011	1500	9.9	0.286

Table 2: Model inputs relating to oiling

Constant	Description	Value
N_{1994}^{oil}	Number of adult females which died as a result of the 1994 oil spill	600
N_{2000}^{oil}	Number of adult females which died as a result of the 2000 oil spill	750
p_{1994}^{oil}	Proportion of chicks which died due to the 1994 oil spill	0.29
p_{2000}^{oil}	Proportion of chicks which died due to the 2000 oil spill	0.38
m_{1994}	Month of the 1994 oil spill	5.7
m_{2000}	Month of the 2000 oil spill	5.9

Table 3: Sardine and anchovy survey biomass in thousands of tons. The May recruit survey biomass is for the entire standard survey area west of Cape Infanta. The November biomass is for the proportion west of Cape Agulhas (strata A–C).

Year	November spawner survey		May recruit survey	
	Sardine	Anchovy	Sardine	Anchovy
1984	48.009	1461.636		
1985	25.457	1014.215	38.265	368.623
1986	238.230	1978.652	50.073	621.089
1987	94.165	1866.430	98.643	721.578
1988	128.043	1289.624	5.223	563.107
1989	198.328	517.293	66.081	173.349
1990	248.855	342.812	31.208	170.083
1991	517.180	1254.359	26.665	528.177
1992	247.756	1036.580	74.822	458.455
1993	480.822	439.121	114.956	481.108
1994	389.730	309.981	72.462	145.336
1995	348.832	468.678	205.149	392.016
1996	257.763	29.748	73.612	74.842
1997	964.835	377.663	396.718	404.620
1998	1082.547	206.586	134.907	453.210
1999	708.029	741.961	235.720	826.090
2000	726.230	1960.122	299.473	2553.502
2001	669.617	2301.999	573.427	1998.427
2002	1184.713	2018.570	616.331	1560.101
2003	1343.118	1181.111	600.667	1434.900
2004	292.522	736.973	40.419	1071.419
2005	75.604	670.730	11.236	560.518
2006	177.885	1027.009	50.394	275.797
2007	57.666	889.676	34.575	1534.523
2008	211.871	1421.593	24.461	1491.847
2009	262.853	2098.253	63.468	1317.059
2010	309.465	354.148	499.986	1687.118

Table 4: Parameter values fixed on input

Parameter	Description	Value
A	Plus group age	5
p_J	Detectability of juveniles relative to adults in the moult count	1.0
a^*	Age of first breeding attempt	4
σ_J	Standard errors of the logarithms of the juvenile proportions	0.1
$\tilde{\sigma}$	Parameter related to variability about adult mortality vs fish abundance	0.1
σ_H	Standard deviation of reproductive success	0.1
H_{max}	Maximum allowed reproductive success	1.8
q_M	Proportion of moulters susceptible to observation	0.9
M_{min}	Minimum allowed mortality rate	0.04

Table 5: List of estimated parameters, prior distributions and Bayesian posterior probability intervals

Parameter	Description	Prior	Mode	5%	Median	95%
$p_{TR,y}$	Re-sighting probabilities	U[0, 1]				
M_{trans}	"transient" mortality of tagged birds	U[0, 1]	0.349	0.268	0.354	0.446
$M_{trans,1994}$		U[0, 1]	0.000	0.000	0.000	0.001
$M_{trans,2000}$		U[0, 1]	0.000	0.000	0.000	0.000
$\ln N_0$	Log of initial population size	U[1, 10]	6.853	6.469	6.853	7.238
λ	Initial population profile parameter	U[0, 3]	0.211	0.176	0.213	0.256
σ_{add}	Additional variance in moult counts	U[0, 1]	0.000	0.000	0.000	0.000
$I_{1989-1991}$	Immigration of three year old birds	U[0, 3000]	561.1	285.7	567.6	852.5
$I_{1992-1994}$		U[0, 3000]	1121.6	690.9	1024.5	1373.1
$I_{1995-1999}$		U[0, 3000]	42.1	9.2	95.9	271.6
M_1	Biomass–mortality relationship	U[0, 0.5]	0.096	0.004	0.010	0.021
M_2	Biomass–mortality relationship	U[0, 0.5]	0.039	0.010	0.069	0.141
M_3	Biomass–mortality relationship	U[0, 0.5]	0.119	0.023	0.148	0.251
X_y	Adult mortality random effects	U[-4.5, 4.5]				
h	Reproductive success relationship	U[0, 1]	0.520			
H_y	Reproductive success	U[0.0001, 0.9999]				
$-\ln L_M$	Moult count likelihood		−42.5			
$-\ln L_J$	Juvenile proportion likelihood		−48.0			
$-\ln L_{T\text{-}R}$	Tag re-sighting likelihood		6113.3			
P_S	Prior on X_y parameters		−6.0			
P_H	Prior on H_y^* parameters		−23.9			
$-\ln P_{post}$	Total negative log posterior		5993.0			

Table 6: Results of the sensitivity analyses in terms of the numbers of observable female adult moulters in 2009 and projected in 2020 for alternative future sardine biomass B_S.

Future $B_S = 0.2$	2009	2020	2020/2009	ratio to base case
Base case	1694	569	0.34	
1. B–M relationship 1	1732	853	0.49	1.47
2. Force tag data fit	2077	883	0.43	1.27
3. $p_J = 0.9$	1682	594	0.35	1.05
4. $a^* = 3$	1714	537	0.31	0.93
5. $a^* = 5$	1678	652	0.39	1.16
6. $\sigma_J = 0.2$	1733	544	0.31	0.93
7. $\tilde{\sigma} = 0.05$	1762	571	0.32	0.96
8. $\tilde{\sigma} = 0.2$	1657	607	0.37	1.09
9. $H_{max} = 1.5$	1676	568	0.34	1.01
10. $q = 0.8$	1694	568	0.34	1.00
11. $q = 1.0$	1695	571	0.34	1.00
Future $B_S = 0.3$				
Base case	1694	2180	1.29	
1. B–M relationship 1	1732	2288	1.32	1.03
2. Force tag data fit	2077	3601	1.73	1.35
3. $p_J = 0.9$	1682	2509	1.49	1.16
4. $a^* = 3$	1714	2116	1.23	0.96
5. $a^* = 5$	1678	2278	1.36	1.06
6. $\sigma_J = 0.2$	1733	2138	1.23	0.96
7. $\tilde{\sigma} = 0.05$	1762	3028	1.72	1.34
8. $\tilde{\sigma} = 0.2$	1657	1232	0.74	0.58
9. $H_{max} = 1.5$	1676	2188	1.31	1.02
10. $q = 0.8$	1694	2170	1.28	1.00
11. $q = 1.0$	1695	2198	1.30	1.01

Figures

Sardine spawners
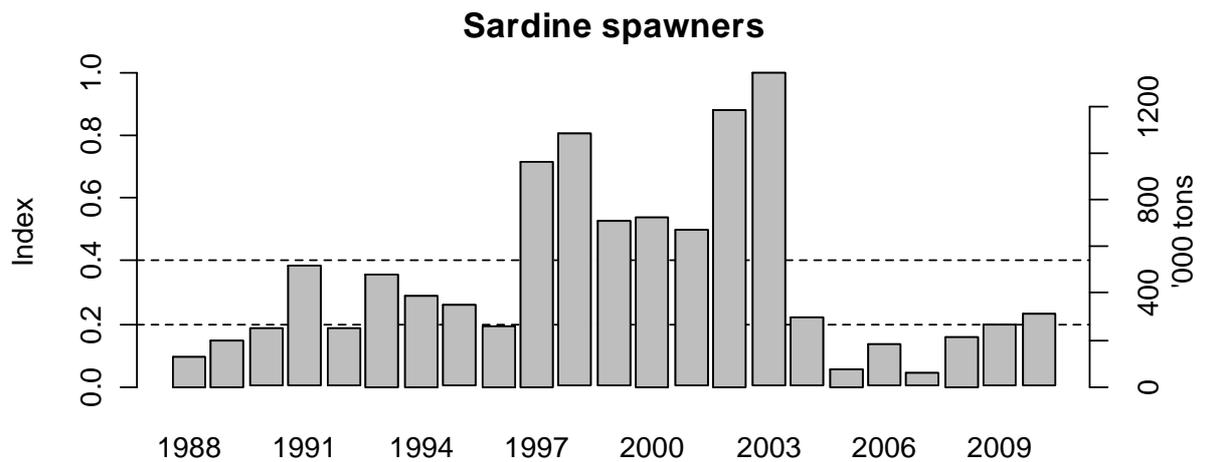
Figure 1: The index for sardine November survey biomass west of Cape Agulhas, which is expressed relative to the maximum in 2003. Dashed horizontal lines at index values 0.2 and 0.4 indicate the range of biomass levels used for the projections illustrated in Figure 11.

Adult mortality/Fish abundance relationship
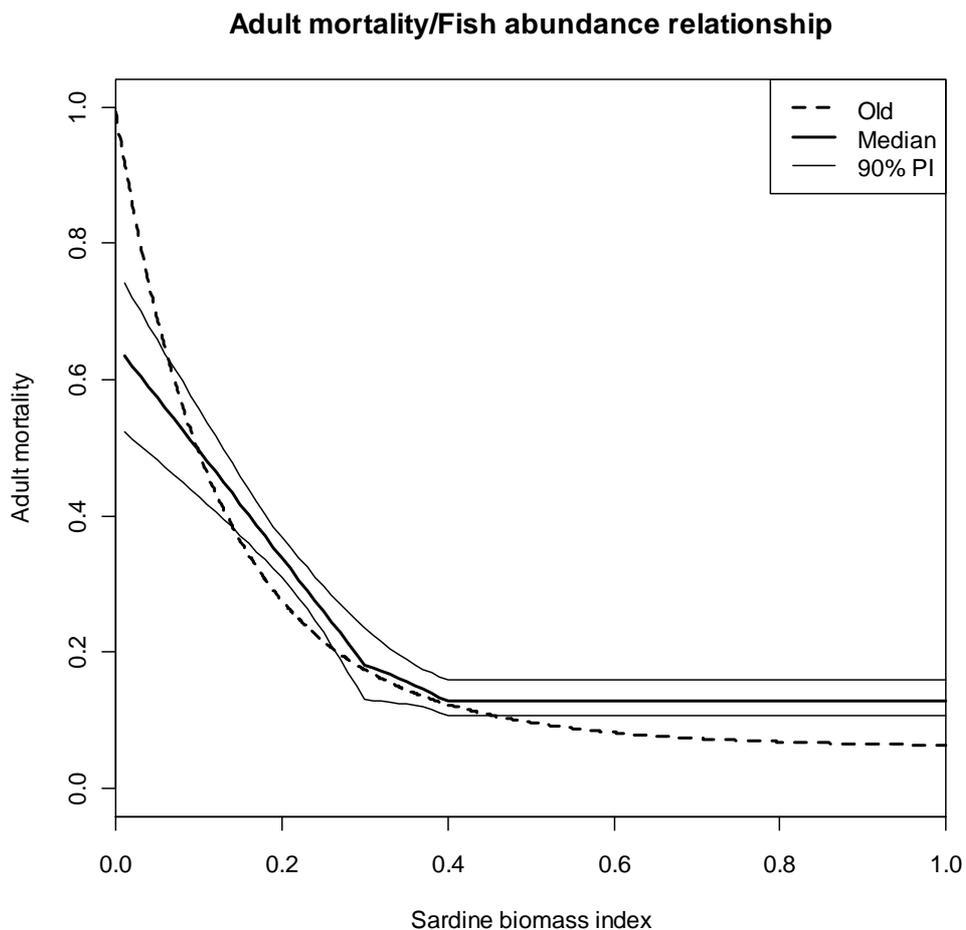
Figure 2: Solid curves show the posterior median and 90% probability interval of the relationship (equation (14)) between penguin adult mortality and the sardine spawner biomass west of Cape Agulhas for the new base case. The dashed curve shows the posterior median of the parametric relationship (equation (11)) which was estimated earlier.

15

Probability of re-sighting

Number of re-sightings
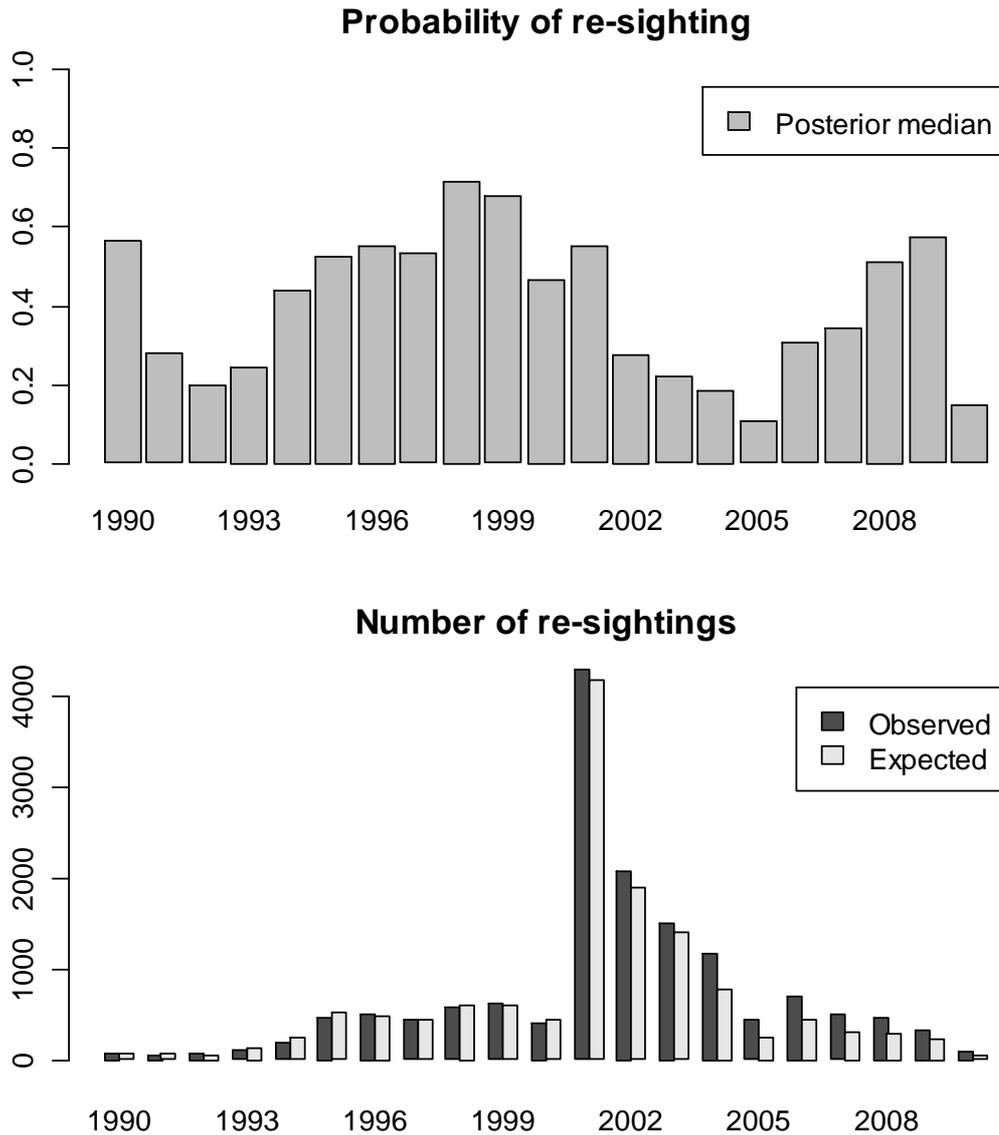
Figure 3: Results of fitting to tag-recapture data for the new base case. Top: annual probability of re-sighting banded birds. Bottom: comparison of observed and expected numbers of banded penguins re-sighted each year for the joint posterior mode.
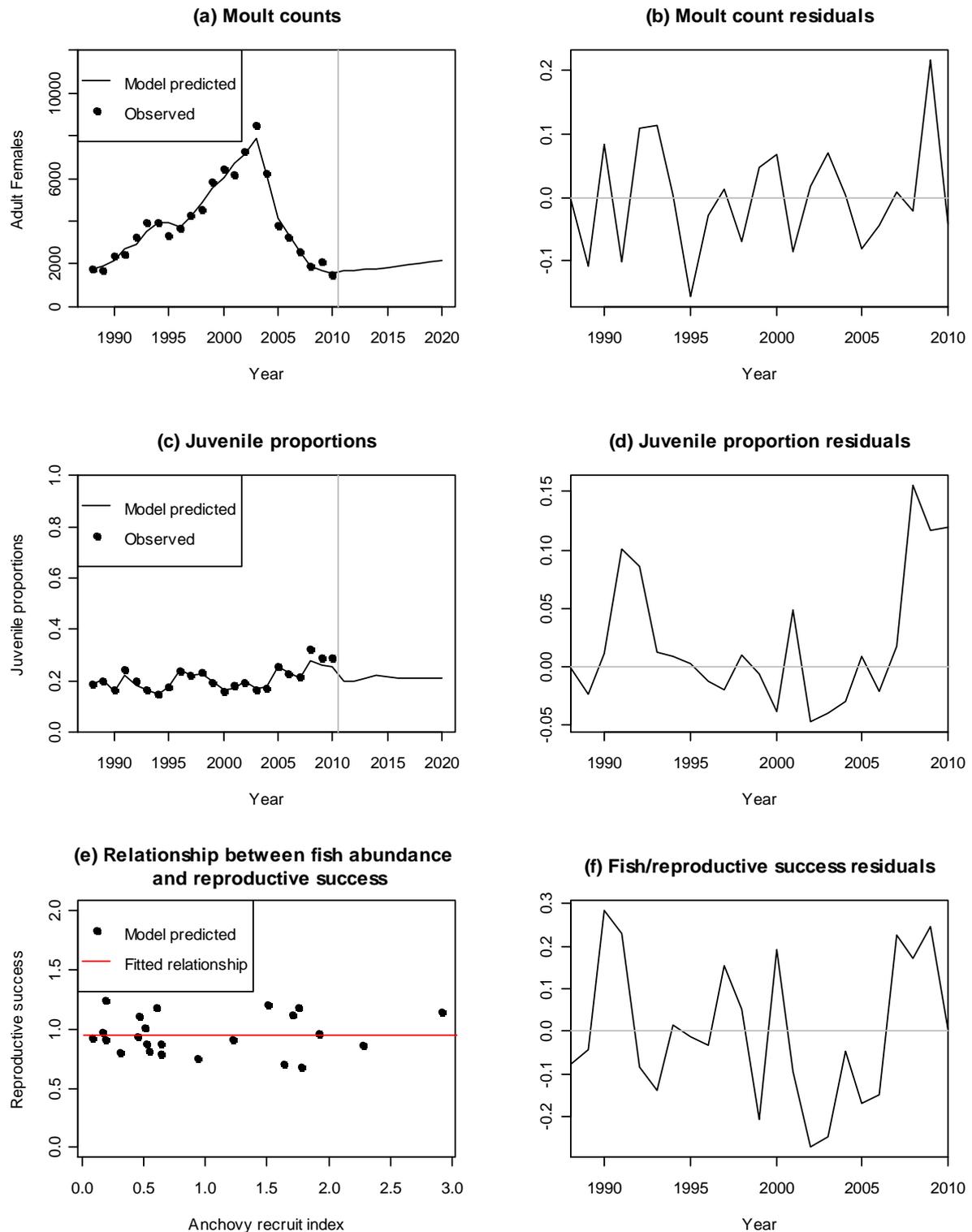
Figure 4: Fits to data/relationships and associated residuals at the joint posterior mode for the new base case (which uses equation (14) for adult penguin natural mortality). Projections beyond 2010 assume $B_S = 0.3$. The residuals in (b) and (d) are the differences between the logarithms of the observations and the model predicted values. The residuals in (f) are the differences between the estimated reproductive success rates and the assumed relationship.
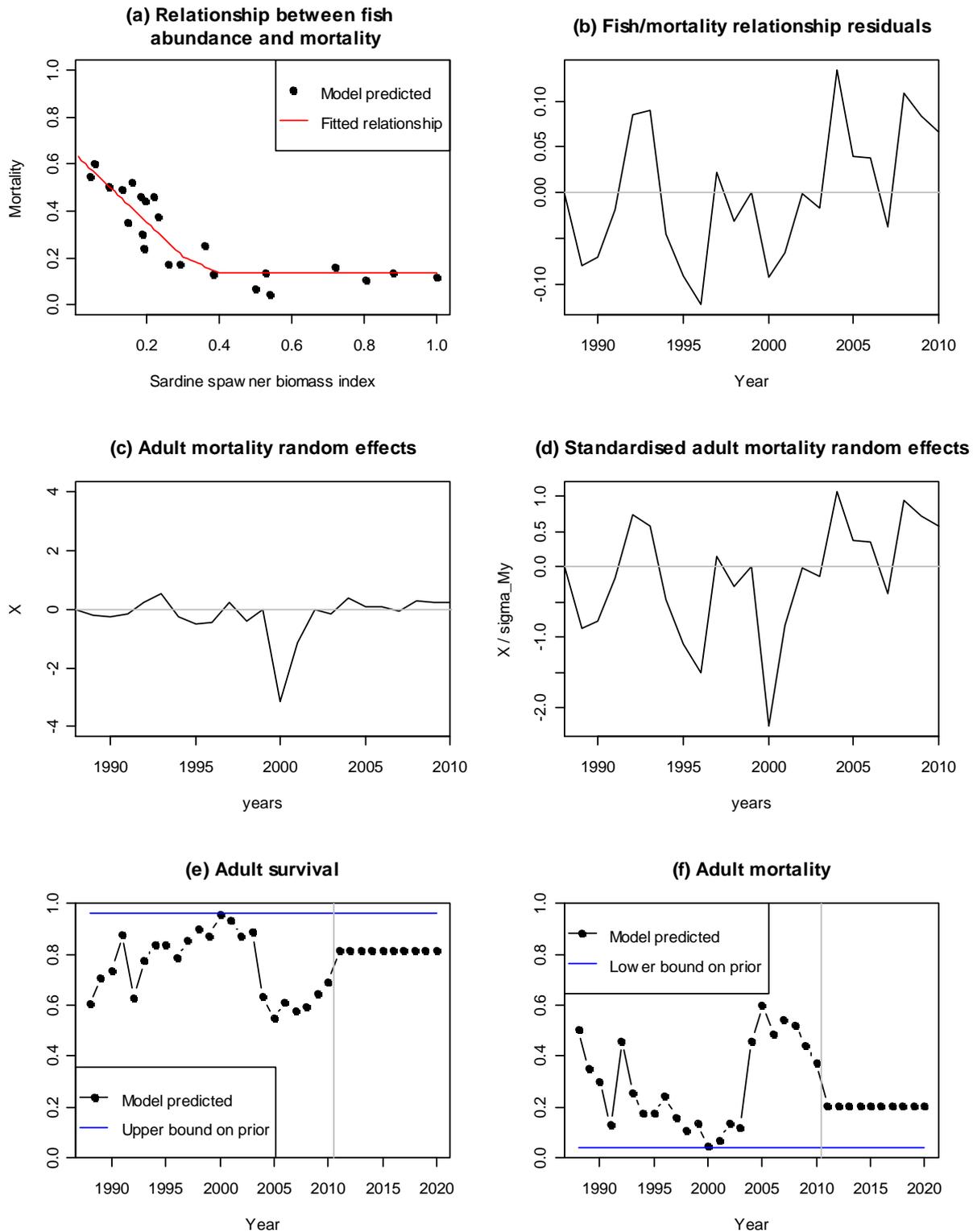
Figure 5: Further fits to data/relationships and associated residuals for the joint posterior mode for the new base case (which uses equation (14) for adult penguin natural mortality). Projections beyond 2010 assume $B_S = 0.3$. Residuals in (b) are the differences between the estimated annual mortality rates and those predicted by the relationship with fish abundance. The random effects X_y are in (c), and are standardised in (d) by dividing by σ_y. Plot (e) shows the time series of the estimates for the adult survival rates, and (f) shows the corresponding mortality rates.
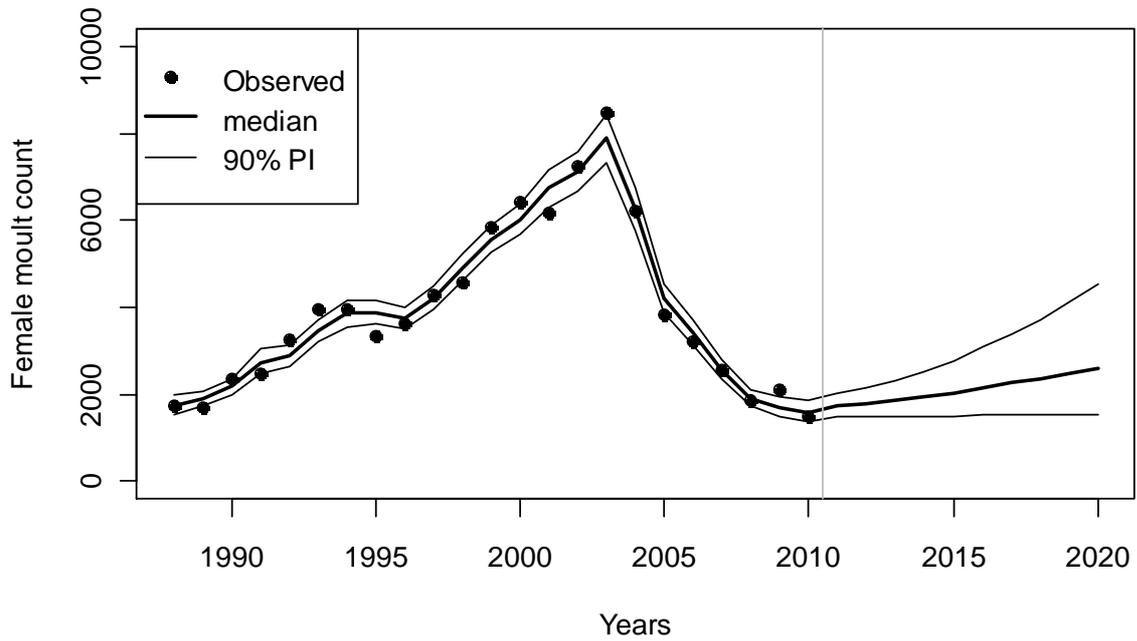
Figure 6: Time series of observed counts of female moulting penguins and the median and 90% probability interval of the Bayesian posterior distribution of the model predicted moult counts for the new base case. The projections assume future sardine biomass west of Agulhas $B_{\mathrm{s}} = 0.3$.
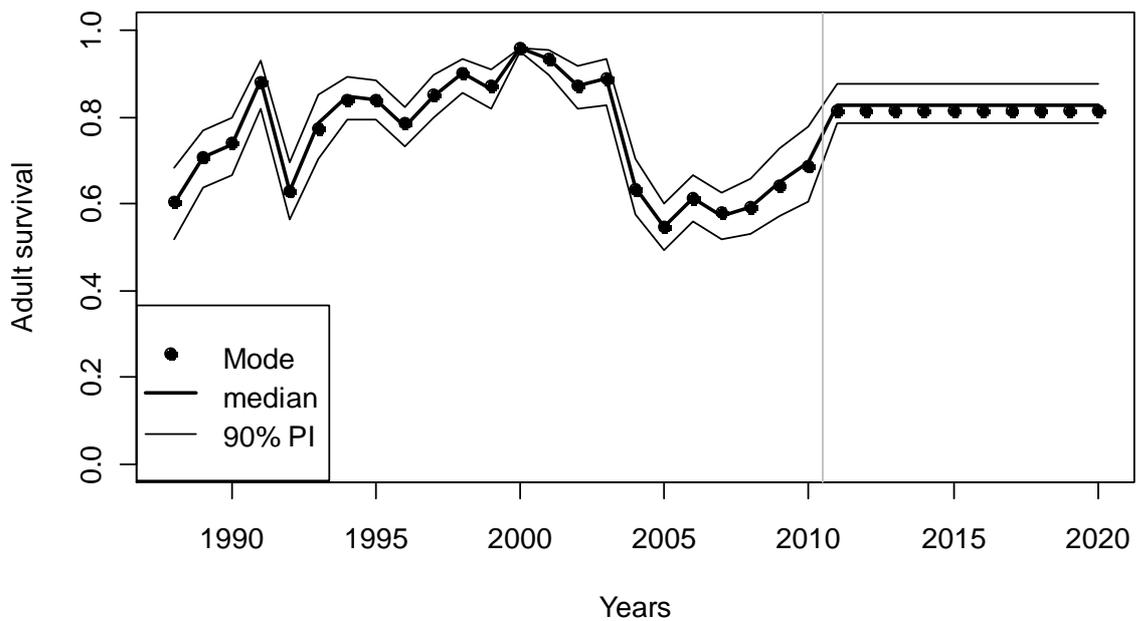
Figure 7: Time series of the adult annual survival rates at the joint posterior mode and the median and 90% probability interval of the Bayesian posterior distribution for the new base case. The projections assume future sardine biomass west of Agulhas $B_{\mathrm{s}} = 0.3$.

Posterior distributions
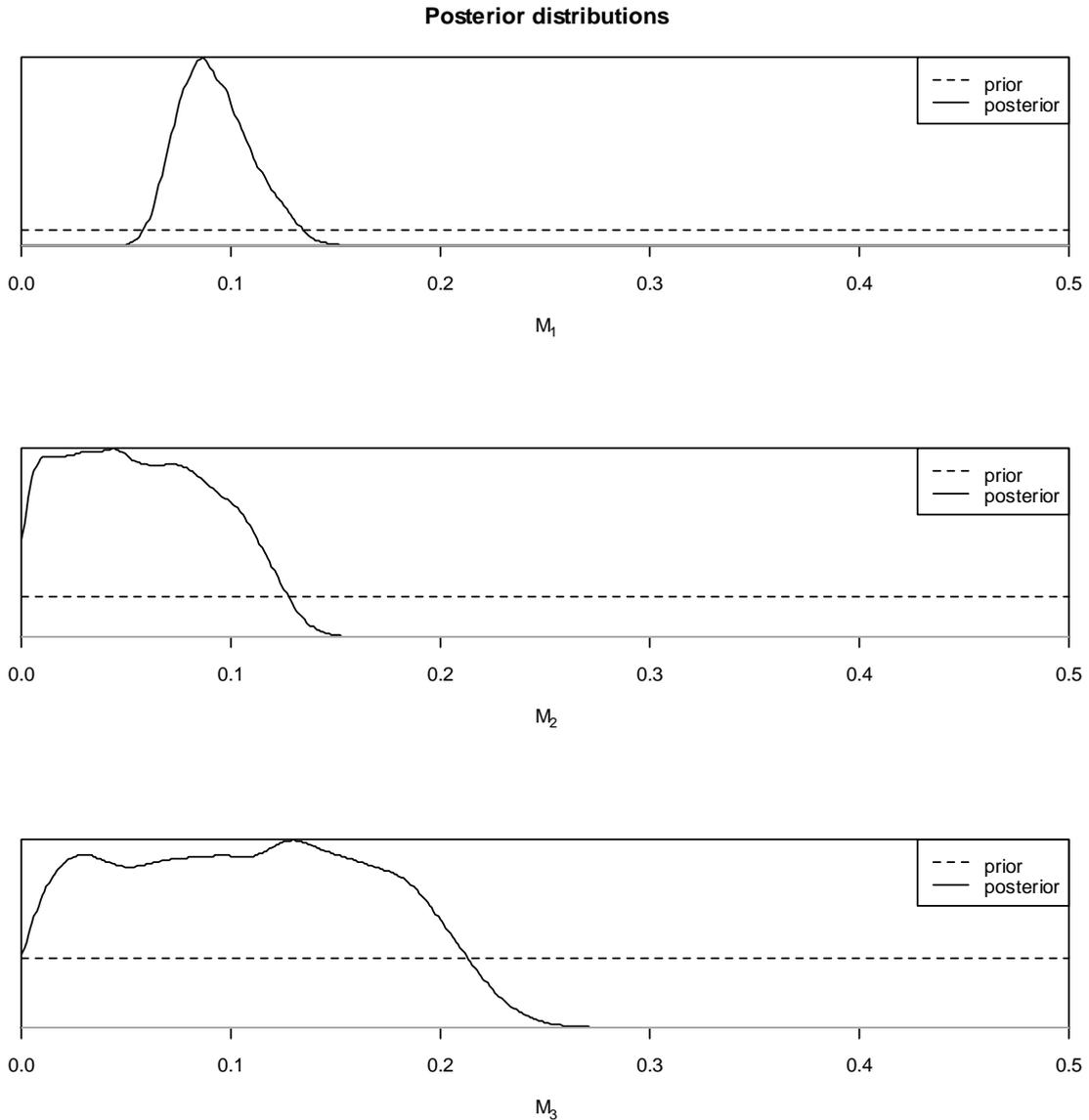
Figure 8: Prior and posterior distributions of the parameters in the mortality–biomass relationship (equation (14)) for the new base case.
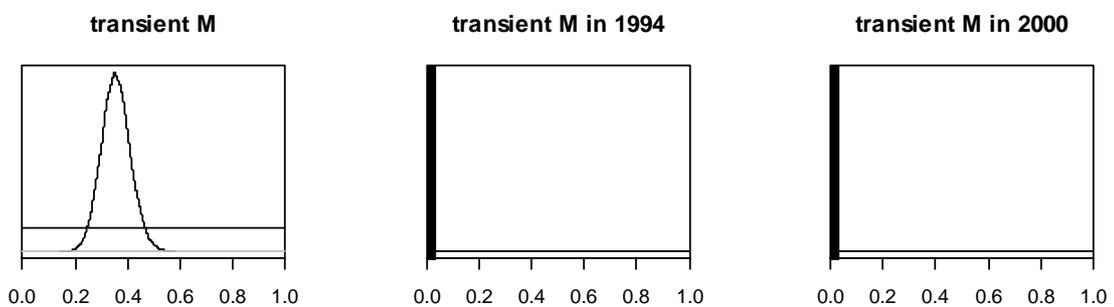
Figure 9: "Transient" mortality comparison of priors and posteriors for the new base case. The "transient M" applies for all years except those with major oil spills (1994 and 2000) from which many penguins were tagged for which separate estimates are made. The solid black bars indicate the posterior to be entirely at transient M=0.

1989-1991	1992-1994	1995-1999
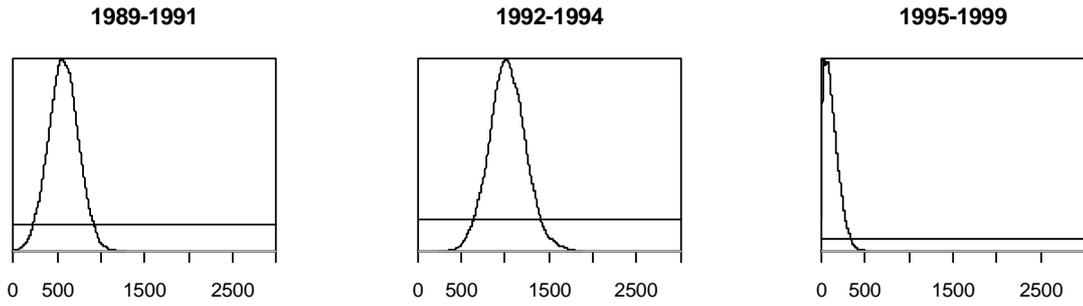
Figure 10: Immigration, taken to be constant over the periods shown: comparison of priors and posteriors for the new base case.
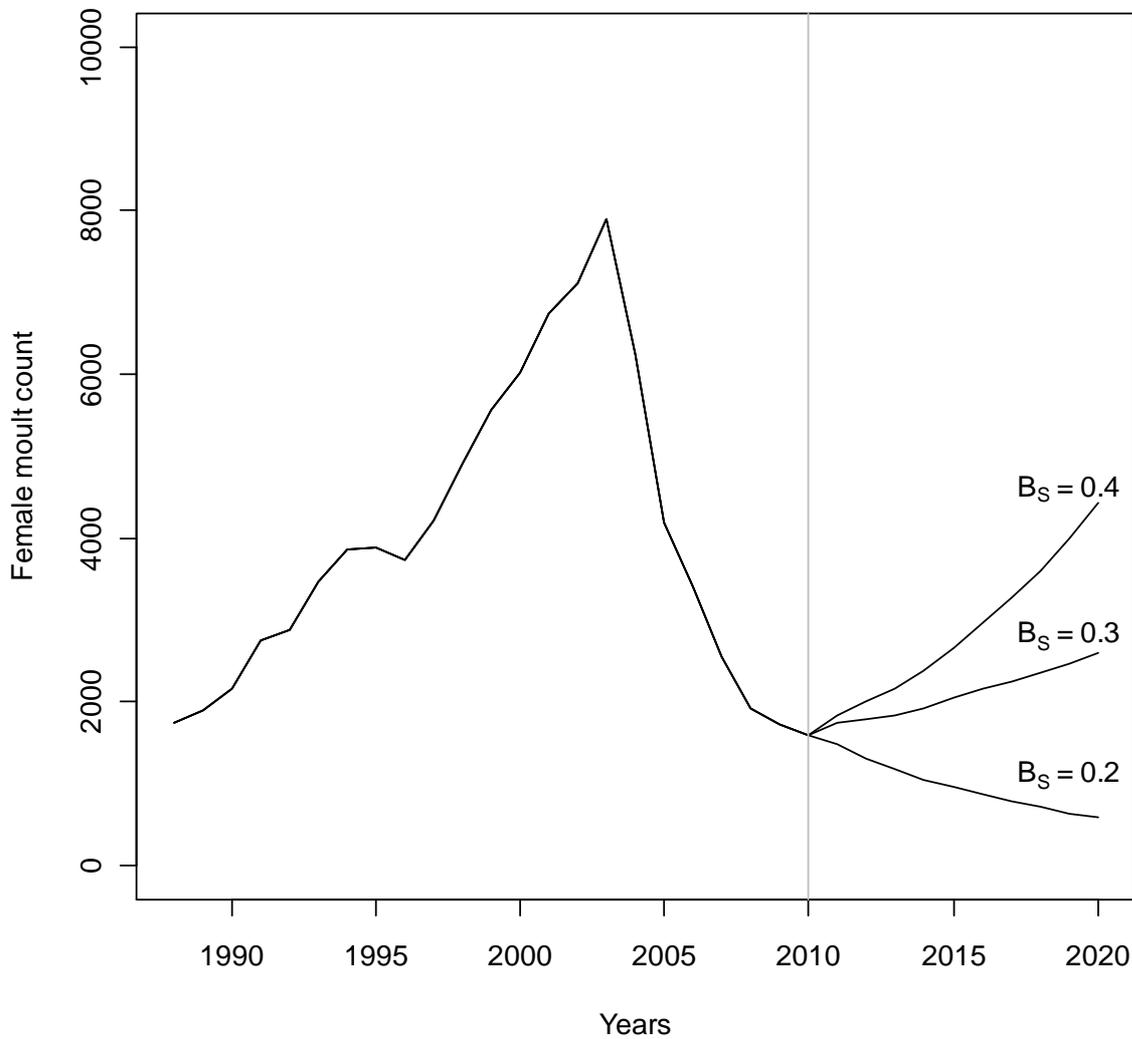
Figure 11: Projections of penguin moult counts from the joint posterior mode for the new base case from 2010 for fixed future sardine spawner biomass levels west of Cape Agulhas for $B_S = 0.2$, $B_S = 0.3$ and $B_S = 0.4$.
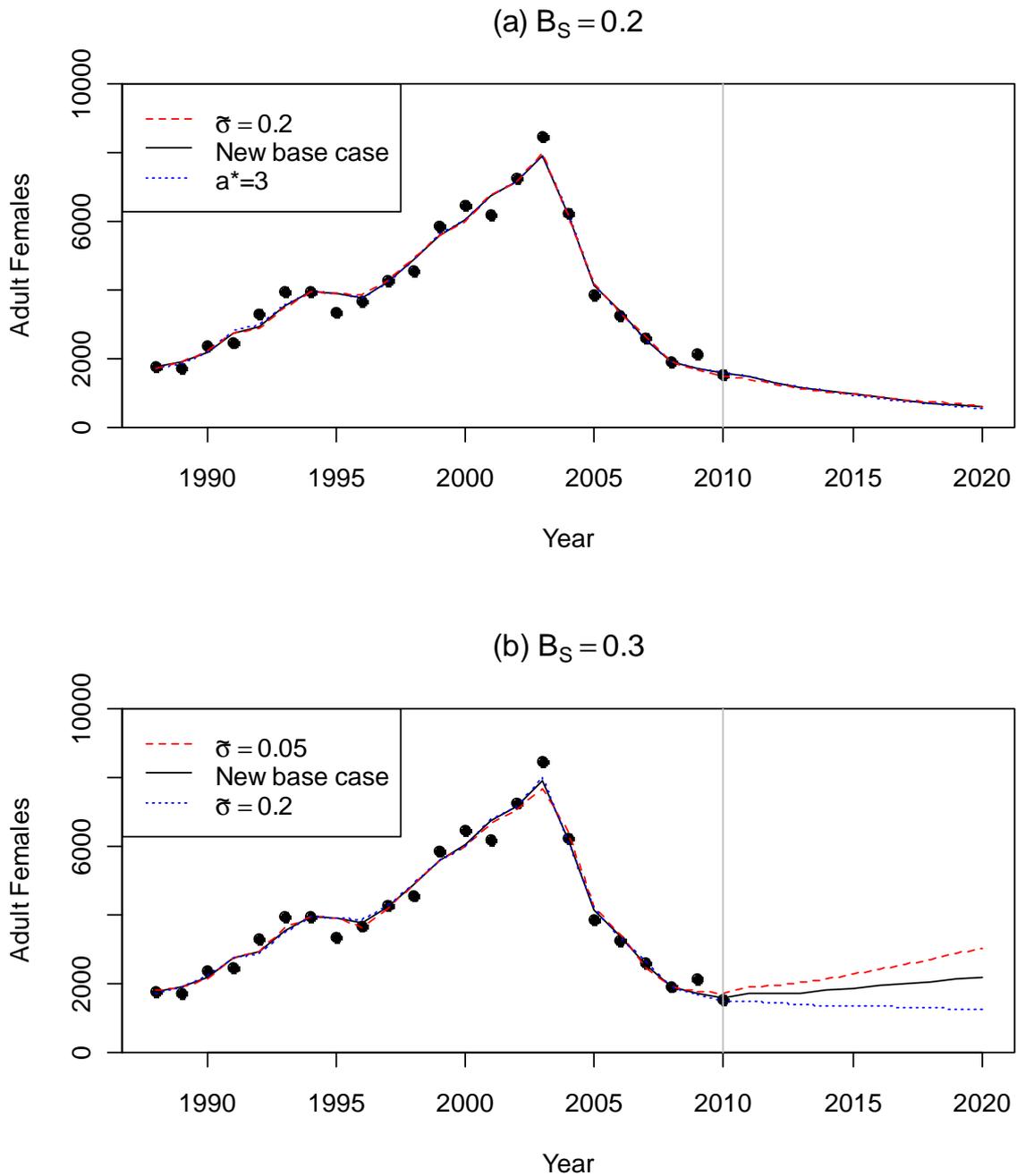
21

(a) $B_S = 0.2$

(b) $B_S = 0.3$

Figure 12: The fits to moult counts and projections of penguin numbers for some sensitivities are shown: (a) at lower future sardine biomass $(B_S = 0.2)$ increasing the value of the variability about the mortality versus biomass relationship $\tilde{\sigma}$ to 0.2 was found to have a small positive influence on projected penguin numbers, while reducing the age at which penguins first attempt breeding to $a^* = 3$ had a small negative effect; (b) at a higher future sardine biomass level $(B_S = 0.3)$, a large positive effect is obtained by reducing the variability about the mortality versus biomass relationship $\tilde{\sigma}$ to 0.05, while increasing $\tilde{\sigma}$ to 0.2 has a negative effect.
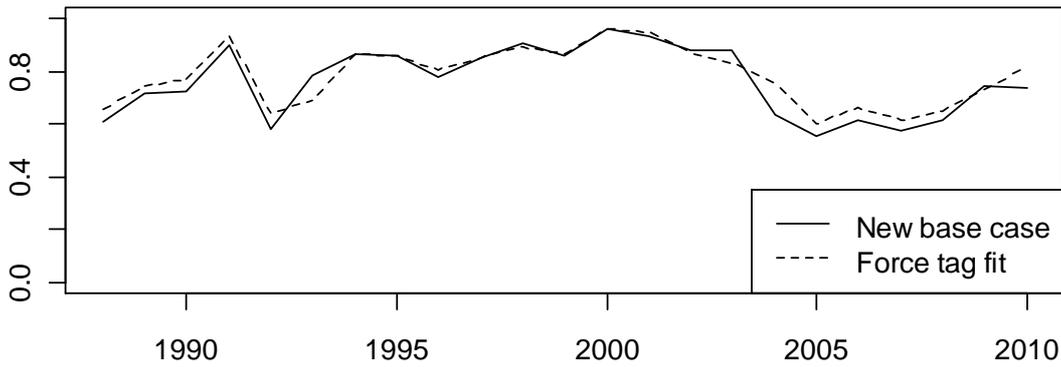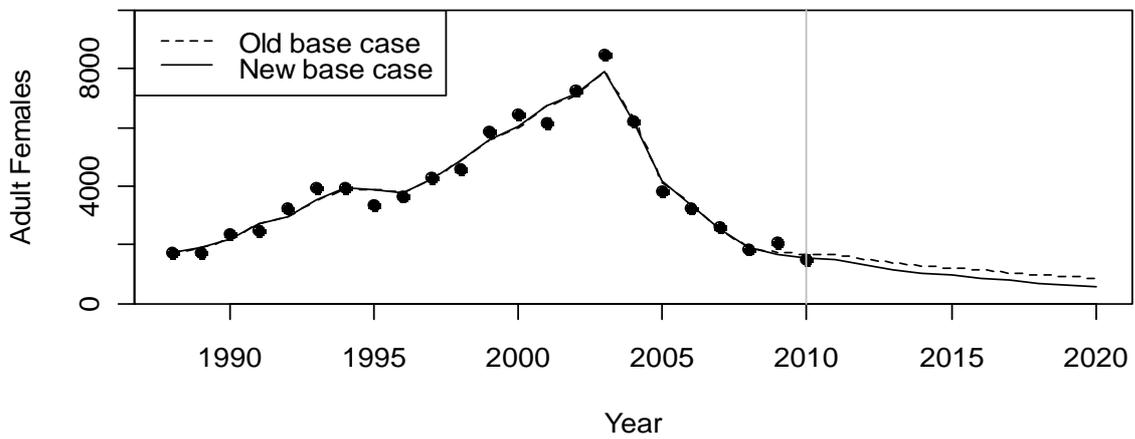
Adult survival

Figure 13: Comparison of adult survival rates for the new base case and the model variation where expected numbers of tag re-sightings are forced to match the observed counts.

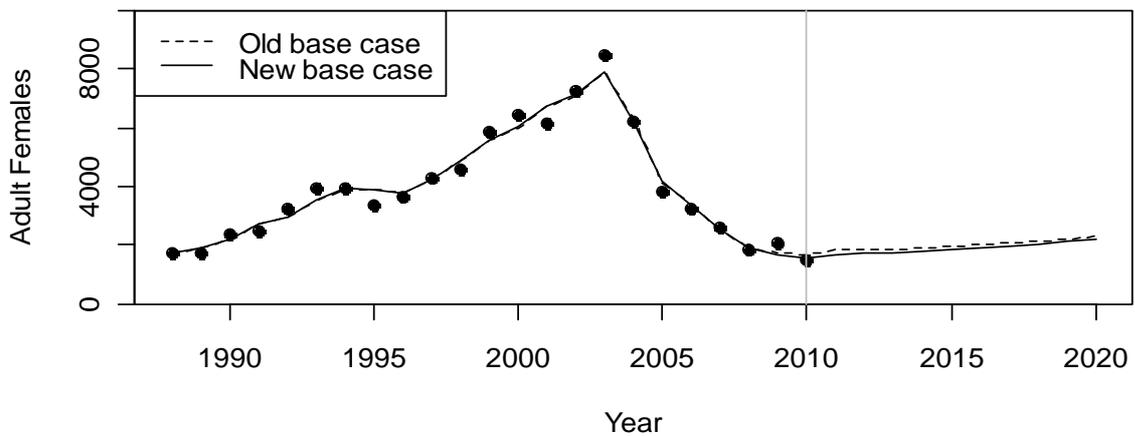(a) $B_S = 0.2$

(b) $B_S = 0.3$

Figure 14: The fits and projections for the old base case (equation (11)) and the new base case (equation (14)) mortality–biomass relationships for future sardine biomass levels $B_S = 0.2$ and $B_S = 0.3$.