

## Updated Results from a Spatial Age-Structured Model of African Penguin populations for use in linking to the pelagic OMP testing process

William Robinson and Éva Plagányi  
Department of Mathematics and Applied Mathematics, UCT

### SUMMARY

Updated results of an earlier penguin model are presented. The model focuses on four colonies 1) Robben Island; 2) Dassen Island; 3) Boulders and 4) Dyer island, although projections will narrow the focus further to the first two of these. Some progress has been made in terms of fitting the model adequately to historic data. However, there are still problems with the fits such that the current model is not yet at the stage where it can be linked to the pelagic OMP (Operational Management Procedure) testing process to take account of the relationship between the breeding success and also natural mortality of African penguins and the abundance of both anchovy *Engraulis encrasicolus* and/or sardine *Sardinops sagax*. Despite extensive investigations, in many cases no overly strong functional relationships between penguin breeding success and pelagic abundance have been found. As a result, functional relationships are being investigated between the natural mortality of African penguins and pelagic abundance, with results suggesting that a relationship exists at Robben Island, but not at Dassen Island. Dassen and Robben Islands are split and modelled separately for these and other reasons.

### INTRODUCTION

This paper provides an updated description of a model of African penguin *Spheniscus demersus* dynamics. The model is of a form that can readily be linked to the pelagic OMP (Operational Management Procedure) testing process to take account of the relationship between the breeding success and also natural mortality of African penguins and the abundance of both anchovy *Engraulis encrasicolus* and/or sardine *Sardinops sagax* (e.g. Crawford *et al.* 2006).

The model is spatial in that several different populations of penguins are represented, and movement between these populations can be modelled. The main focus of the model is on Dassen and Robben Islands, which are here modelled separately (but with some parameter estimates assumed common to both). The third population is Dyer Island because it is the next largest, recent declines in the population there are of concern and it is considered an important breeding site for penguins given the eastward shift of sardines. The fourth population is Boulders. Although relatively small, this colony was considered important to include because of its position, its role as the focus of several other studies and because penguins are known to have moved from Dyer island to Boulders, Robben and Dassen, and that it is useful to quantify to what extent movement of birds away from Dyer island could account for observed declines at Dyer and increases at these other colonies.

The Algoa Bay colony is not assumed linked to the above colonies in the current model version because westward migration is believed to be negligible.

### **Model dimensions**

The PWG agreed that for the purposes of coupling penguin and pelagic fish models, the south-western area should be disaggregated into the following areas:

- i) Cape Columbine to Cape Agulhas
- ii) Cape Agulhas to Algoa Bay
- iii) Algoa Bay to Port Alfred

A summary of all the data inputs are provided in the accompanying document MCM/2008/SWG-PEL/21a. The model time step is one year and hence average trends are modelled. Penguins in each sub-area are modelled from 1987 until 2008.

During the current phase of the model development, only the Western Cape penguin colonies are modelled, corresponding to area i). It was decided to focus on this area because the data series are the most complete, and the majority of the birds are found there. The sardine and anchovy models consist of all THREE areas, with model-predicted biomass in i) being used only when trying to find a functional relationship with the penguin results. The model described here has four sub-areas: 1) Robben Island; 2) Dassen Island; 3) Boulders and 4) Dyer island. Penguins in all of these colonies are assumed to depend on the same prey biomass, namely the west of Cape Agulhas portion of the total pelagic biomass.

The model is coded in AD Model Builder which permits rapid and efficient minimization.

A full description of the model is given in Appendix A, and relationships with food availability are described in Appendix B.

## **MODEL DEVELOPMENT**

The model results are presented in three stages:

- I.) A constant adult survival rate is estimated in the Initial Stage I Model. As found previously, this does not result in an adequate fit to the data (Fig. 1), particularly in terms of failing to adequately represent recent observed declines in the abundance indices. This

is despite modelling breeding success (chiefly in the form of fledging success) as a function of sardine and/or anchovy abundance. This suggests that the relationship between breeding success and pelagic abundance on its own is insufficient to adequately explain the observed trends in penguin abundance.

- II.) An over-parameterised model (hereafter termed the Over-parameterised Stage II Model) is developed which fits the moult count data for Robben Island and Dassen Island exactly (Fig. 2) through estimating residuals that reflect annual changes in survival/mortality of penguins (see Appendix A for technical details). Although it is possible to obtain an excellent fit to the data in this way, a model of this form cannot be projected into the future as a predictive relationship between survival and pelagic abundance is needed. However, the Stage II model is used to examine the nature and existence of such a relationship by plotting the mortality/survival residuals as a function of pelagic abundance (Fig. 5). This suggested a positive linear relationship between penguin survival and pelagic abundance at Robben Island ( $r=0.67$ ,  $n=21$ ,  $p<0.001$ ), but no significant relationship was evident at Dassen Island ( $r=0.34$ ,  $n=21$ ,  $p>0.05$ ).
- III.) The final current model (hereafter referred to as the final current Stage III Model) is re-run using as an input the functional relationship for Robben Island obtained as above between penguin survival and pelagic abundance, i.e. the residuals are not estimated in this model version, but penguin survival at Robben Island varies annually as a function of pelagic abundance. Given that this relationship is fairly weak and has a lot of associated noise, it is not surprising that the model fit worsens substantially (Fig. 6) compared to the Stage II) described above. The associated noise in the functional relationships investigated introduces an added problem in terms of using them for predictive purposes. Note that the Dassen Island penguin survival rate is a constant given the lack of a clear functional relationship for this colony.

## RESULTS AND DISCUSSION

There has been insufficient time to present results for all scenarios here; hence only selected illustrative examples are shown in detail in Figs. 1-6.

The simulations shown assume that first-time breeders move from Dyer to Boulders (to match the Boulders population data) and to Dassen and Robben. The pelagic abundance series used is the combined sardine and anchovy west of Cape Agulhas (see data document).

A few points of interest in the results:

- The results of the revised model version are consistent with previous results presented suggesting that the average and maximum of published survival estimates are too low to sustain populations, even when considering movement from Dyer to Robben/Dassen and Boulders (Table 1).
- The effect of modelling adult survival as a random effect is shown in Fig. 2. In the example shown, the  $\sigma_n$  value was set relatively high (0.6 – Table 1). A random effects term will similarly be added to the juvenile survival rate in further investigations – the main use of the approach is to study the trends in the residuals.
- Fig. 3 summarises three model relationships (mean chicks fledged, the proportion of mature birds that breed each year and the proportion of birds of ages 3 and 4 that first mature each year) that are determined as functions of the pelagic fish abundance input series.
- Note that the juvenile survival rate is modelled as density dependent, and hence it decreases over the last few years following the observed population increases. A low level of density dependence was set in the example shown.
- Fig. 4 shows modelled numbers of birds that moved to Boulders, Dassen and Robben in the illustrative simulation. Given the initial poor fit to the Dyer Island proportion of juveniles data (which show a very low proportion of juveniles), we modified the movement model to allow only the movement of one-year-old birds. Earlier model results (not repeated here) showed that if movement from Algoa Bay westwards was assumed for years when pelagic recruit abundance was high, this could both explain the observed decrease in penguin breeding counts at Algoa Bay, and assist in explaining the observed increases at Dassen and Robben island without having to invoke unrealistically high survival or fecundity estimates. However, task group members noted that there were few observations of penguins emigrating westwards, given the predominant conclusion from movement studies of a counter clockwise movement of birds. It is worth bearing in mind though that the cited studies were conducted prior to the very large increases in pelagic abundance on the west coast and hence further discussion is needed as to the feasibility of the hypothesis that birds moved westwards to exploit the high food availability.
- The fit of the model to the data when random effects are replaced by a fitted relationship to food availability is still unsatisfactory (Fig. 6) for Dassen Island in particular (see also comments in accompanying data document).

## Conclusions

The model developed here has proved a useful tool for exploring various hypotheses and providing a dynamic picture of penguin dynamics. The model has been built up from a very simple form to a more complex form that permits simulating movement of birds between colonies. The simplest form of the model integrates current understanding of penguin population dynamics to test whether the various parameter estimates are compatible with the data. These investigations suggest that there may be problems with both the data used in and assumptions of the model.

For most colonies, the trends in penguin numbers can only be explained if immigration to and emigration from colonies is assumed. The model as described here is still in the process of development and hence results should be viewed as preliminary only.

Note that the Stage III final current model version includes two different functional relationships:

- a. for Robben island only, that between adult penguin survival and pelagic abundance; and
- b. a different functional relationship for each of Robben and Dassen Islands between breeding success and pelagic abundance.

These functional relationships will be used when evaluating predicted future penguin abundance under alternative harvest strategies (OMPs) (see MCM/2008/SWG-PEL/21c). In each case a variety of relationships have been tested, using either sardine, anchovy or both, using a 0, -1 or -2 year time lag, and using total fish biomass or recruit abundance. In all cases pelagic abundance refers to the west of Agulhas component of the total abundance (see MCM/2008/SWG-PEL/21a). However, the penguin model as described here is not yet capable of adequately fitting all available indices when using the derived functional relationships, and hence further work is needed before it can reliably be used to assess the risk to penguins populations under the new pelagic OMP.

## ACKNOWLEDGEMENTS

We are grateful for data from Rob Crawford and Les Underhill, as well as inputs from both them and the rest of penguin task group. Thanks to Janet Coetzee and Carryn de Moor for providing pelagic fish abundance estimates.

**LITERATURE CITED**

- Crawford, R.J.M., Shannon, L.J., Whittington, P.A., 1999. Population dynamics of the African penguin *Spheniscus demersus* at Robben Island, South Africa. *Marine Ornithology* 27, 139-147.
- Plagányi, É.E. and D.S. Butterworth. 2006. A spatial multi-species operating model (SMOM) of krill-predator interactions in small-scale management units in the Scotia Sea. Workshop document presented to WG-EMM subgroup of CCAMLR (Commission for the Conservation of Antarctic Marine Living Resources), WG-EMM-06/12. 28 pp.
- Underhill LG, Bartlett PA, Baumann L, Crawford RJM, Dyer BM, Gildenhuys A, Nel DC, Oatley TB, Thornton M, Upfold L, Williams AJ, Whittington PA, Wolfaardt AC (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.
- Underhill, L. G., Crawford, R. J. M., Wolfaardt, A. C., Whittington, P. A., Dyer, B. M., Leshoro, T. M., Ruthenberg, M., Upfold, L., and Visagie, J. 2006. Regionally coherent trends in colonies of African penguins *Spheniscus demersus* in the Western Cape, South Africa, 1987–2005. *African Journal of Marine Science* 28: 697-704.

Table 1. Summary of example results from the Over-parameterised Stage II model including all Western Cape areas and assuming movement from Dyer to other colonies.

Model	Fit using separate survival residuals for Robben and Dassen				
<u>Parameter estimates</u>					
<i>S</i>	0.82				
<i>S<sub>j</sub>*</i>	0.80				
<i>Emig</i>	0.71				
<i>Imm_Bou</i>	0.09				
<i>Imm_Rob</i>	0.85				
<i>h_Rob</i>	0.85				
<i>h_Das</i>	0.98				
<i>N<sub>0,2+</sub> (Rob)</i>	983				
<i>N<sub>0,2+</sub> (Das)</i>	11660				
<i>N<sub>0,2+</sub> (Dye)</i>	9518				
<u>Likelihood contributions</u>					
<i>-lnL Rob moult</i>	-83.247				
<i>-lnL Rob breed</i>	-20.949				
<i>-lnL Rob prop</i>	-31.149				
<i>-lnL Das moult</i>	-331.905				
<i>-lnL Das breed</i>	-24.254				
<i>-lnL Das prop</i>	-22.476				
<i>-lnL Dye moult</i>	-5.024				
<i>-lnL Dye breed</i>	-14.679				
<i>-lnL Dye prop</i>	-12.817				
<i>-lnL Bou breed</i>	1.567				
<i>-lnL Rob fledge</i>	-5.136				
<i>-lnL Das fledge</i>	-4.643				
<i>S<sub>η</sub> penalty</i>	41.621				
<b><i>-lnL TOTAL</i></b>	<b>-513.090</b>				
<u>Model version</u>					
	<u>Age 1</u>	<u>Age 2</u>	<u>Age 3</u>	<u>Age 4</u>	<u>Age 5</u>
<i>Proportion mature at age</i>	0	0	1.0	1.0	1.0
<i>Pelagic biomass series</i>	Combined anchovy and sardine west of Cape Aguhlas				
<i>Pelagic recruits series</i>	Combined anchovy and sardine recruits				
<i>σ<sub>η</sub></i>	0.6				
	<u>Robben</u>	<u>Dassen</u>	<u>Dyer</u>	<u>Boulders</u>	
<i>K factor</i>	10	10	10	5	
<i>H<sub>max</sub></i>	0.97	1.3717	1.0	1.0	
	<u>1994</u>	<u>2000</u>	<u>prop fledglings</u>		
<i>No's die (oiling)</i>	2500	1000	0.5		

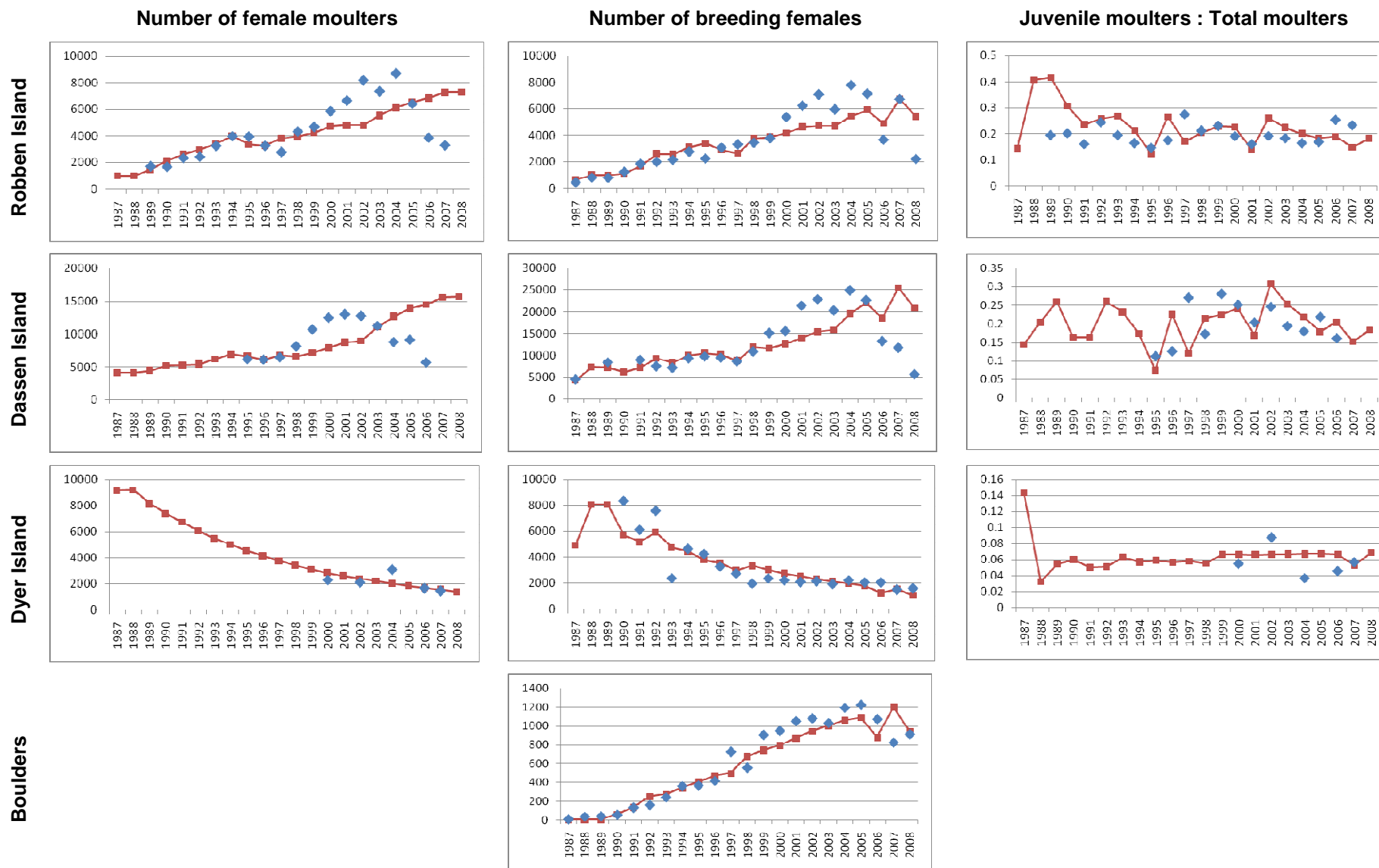


Fig. 1. Illustrative results from the Initial model with a constant survival rate showing model-predicted trajectories of the numbers of female moulting penguins (left panel), number of breeding pairs (centre), and proportion juveniles (right) at Robben (top panel), Dassen (2nd from top); Dyer island (3rd from top), and Boulders. Observed data are shown as diamond points not joined by a line.



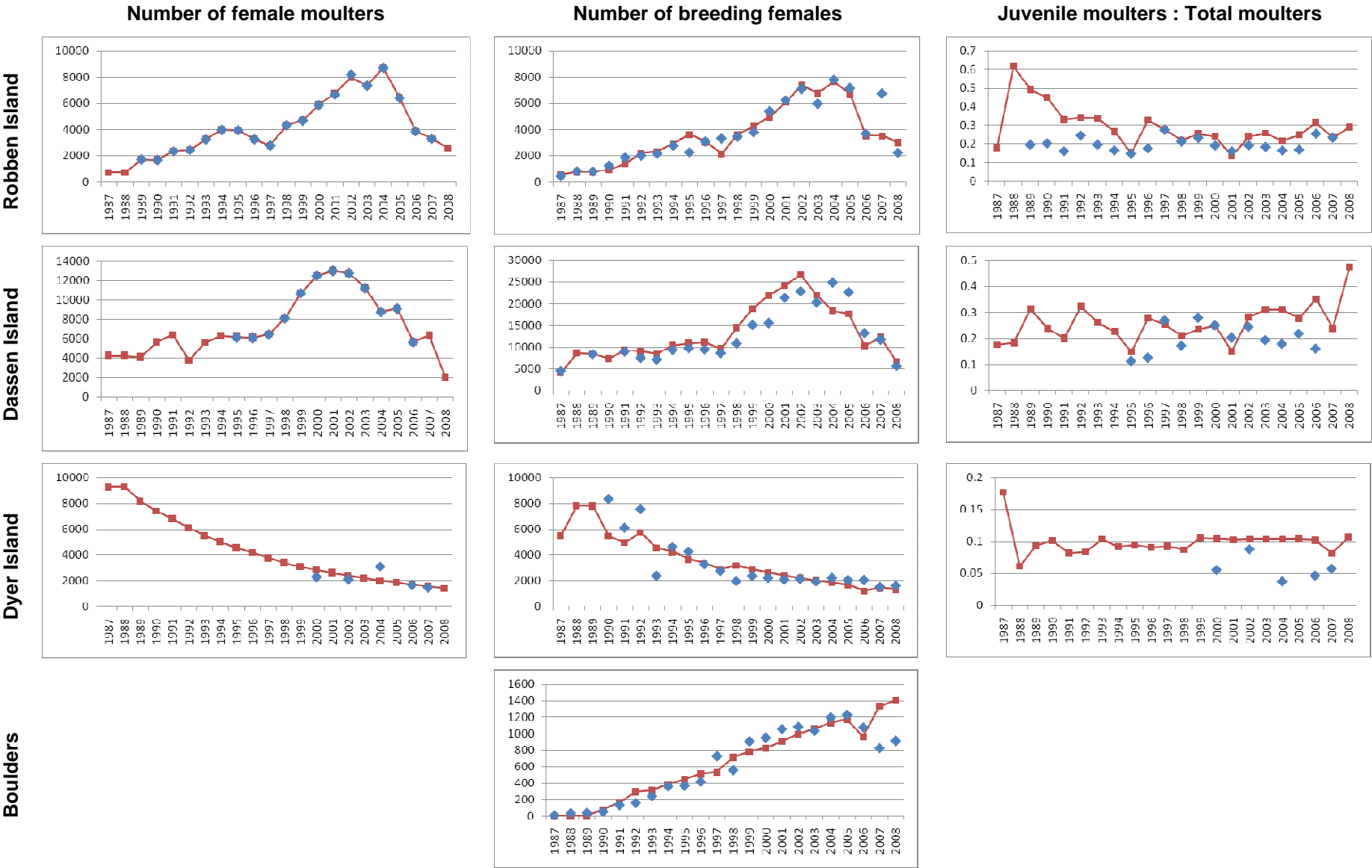


Fig. 2. Illustrative results from Stage II over-parameterised model. Model-predicted trajectories of the numbers of female moulting penguins (left panel), number of breeding pairs (centre), and proportion juveniles (right) at Robben (top panel), Dassen (2nd from top); Dyer island (3rd from top), and Boulders. Observed data are shown as diamond points not joined by a line.

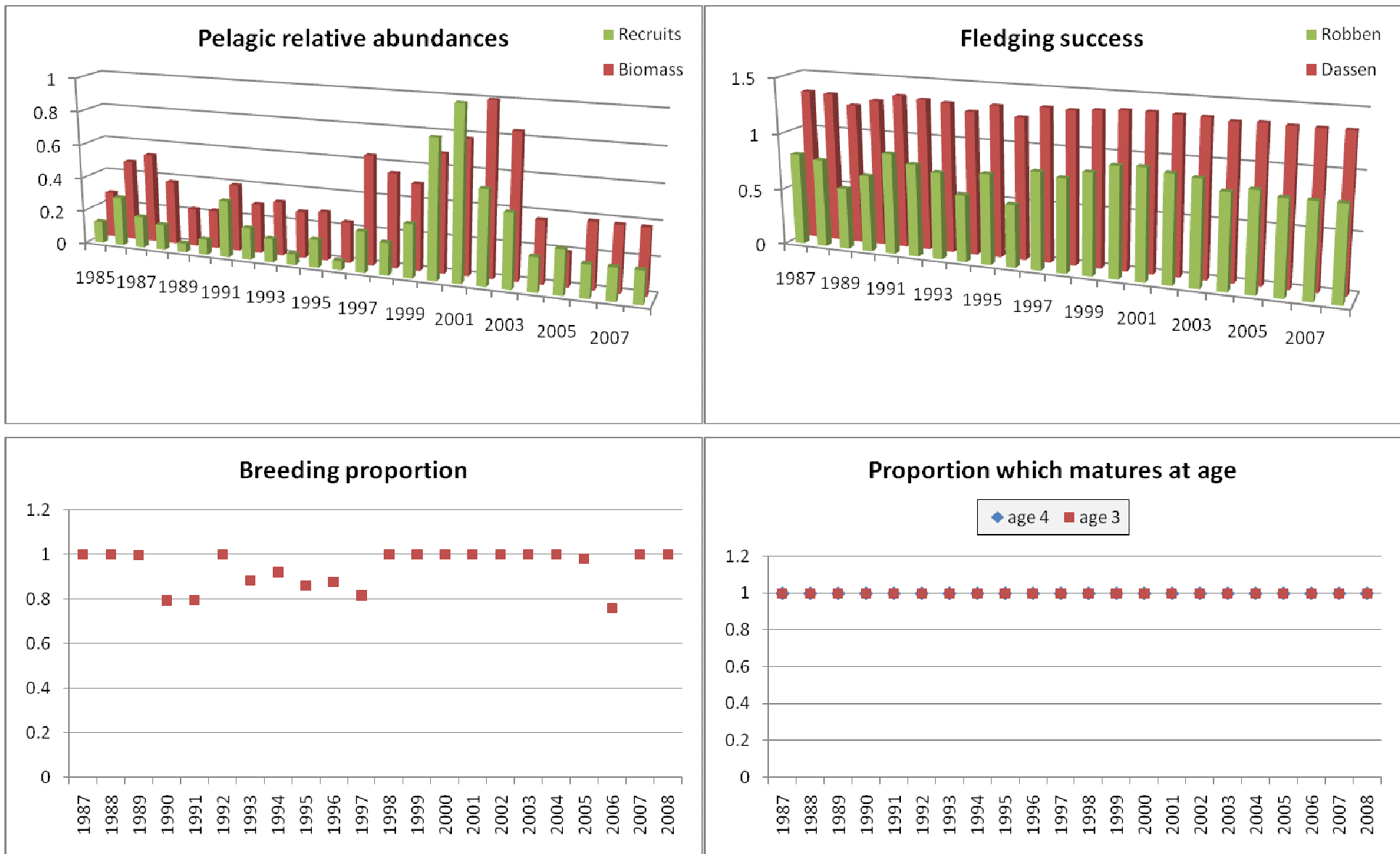


Fig. 3. Illustrative results from Stage II over-parameterised model showing clockwise from top left figure: a) the time series of combined anchovy and sardine recruit and biomass estimates used as model inputs , b) model-estimated fledging success, c) the proportion of mature birds that breed, and d) the proportion of birds that mature at ages as shown.

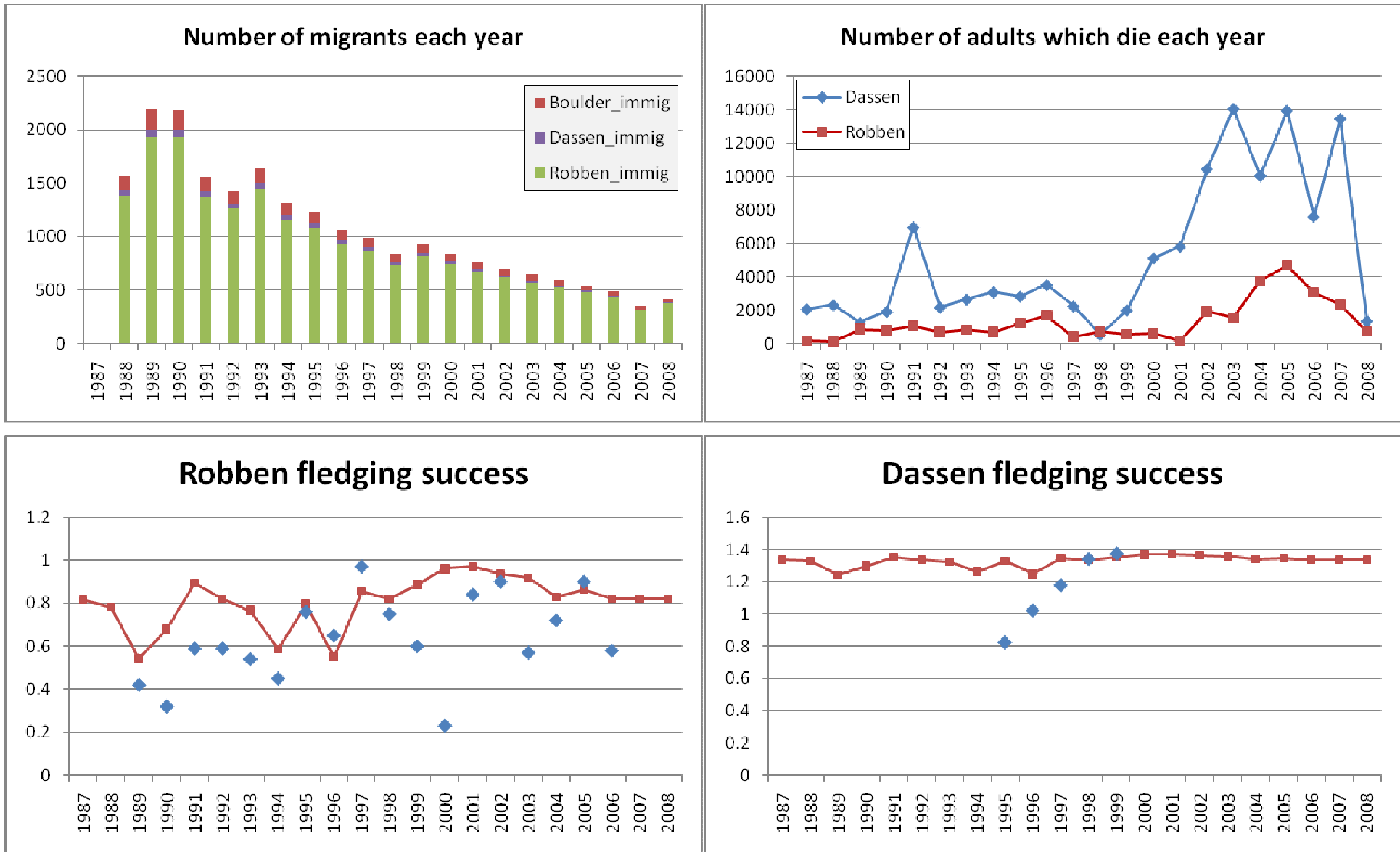


Fig. 4. Illustrative results from Stage II over-parameterised model. Immigration from Dyer Island, the number of adults which die each year, and fledging success at Robben and Dassen.

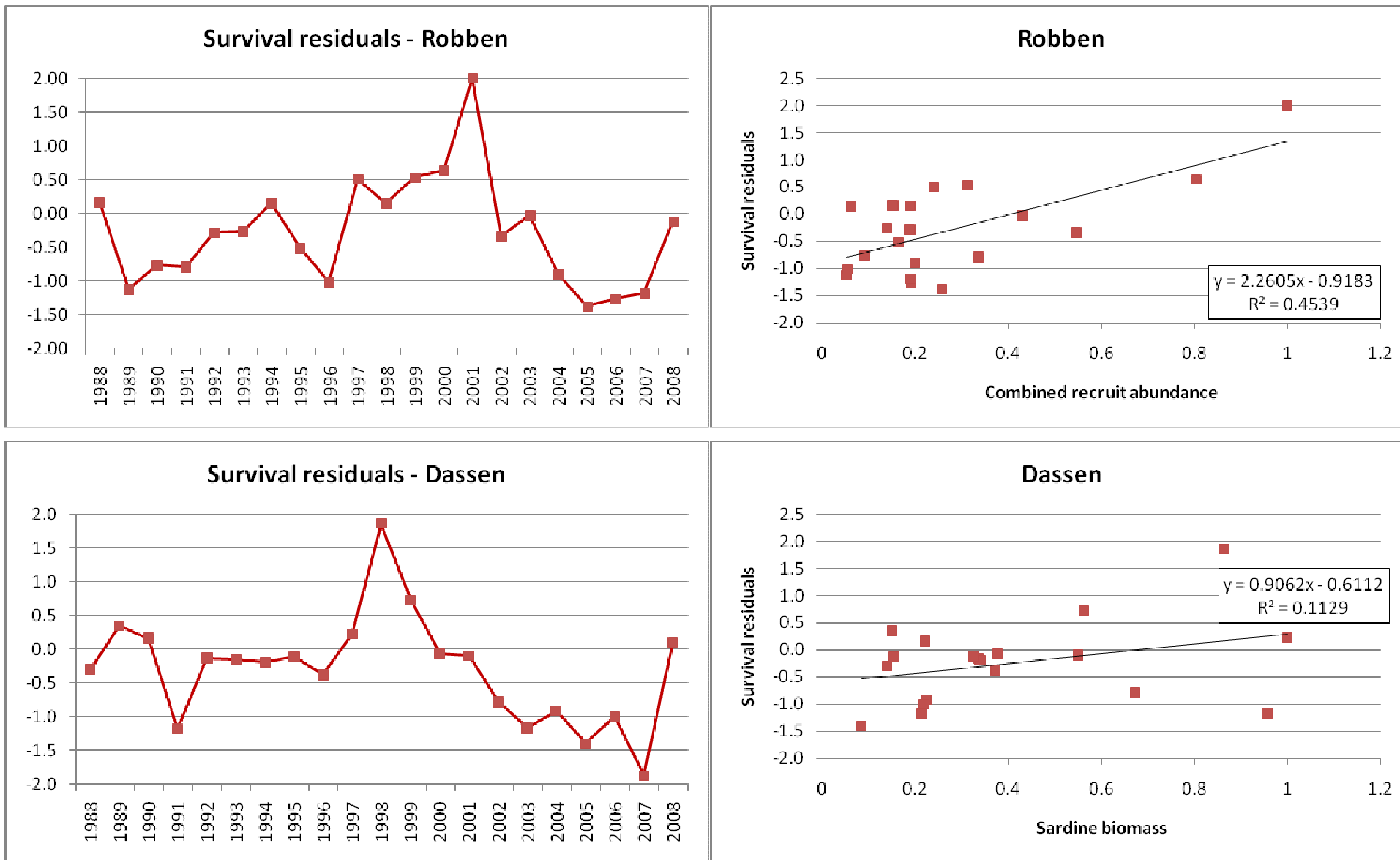


Fig. 5. Illustrative results from Stage II over-parameterised model. The left panel shows the model-estimated survival residuals and shown on the right are the best correlations with pelagic series. The survival residual-pelagic relationships thus obtained are used to define a functional relationship between penguin survival and pelagic abundance for use in later model versions.

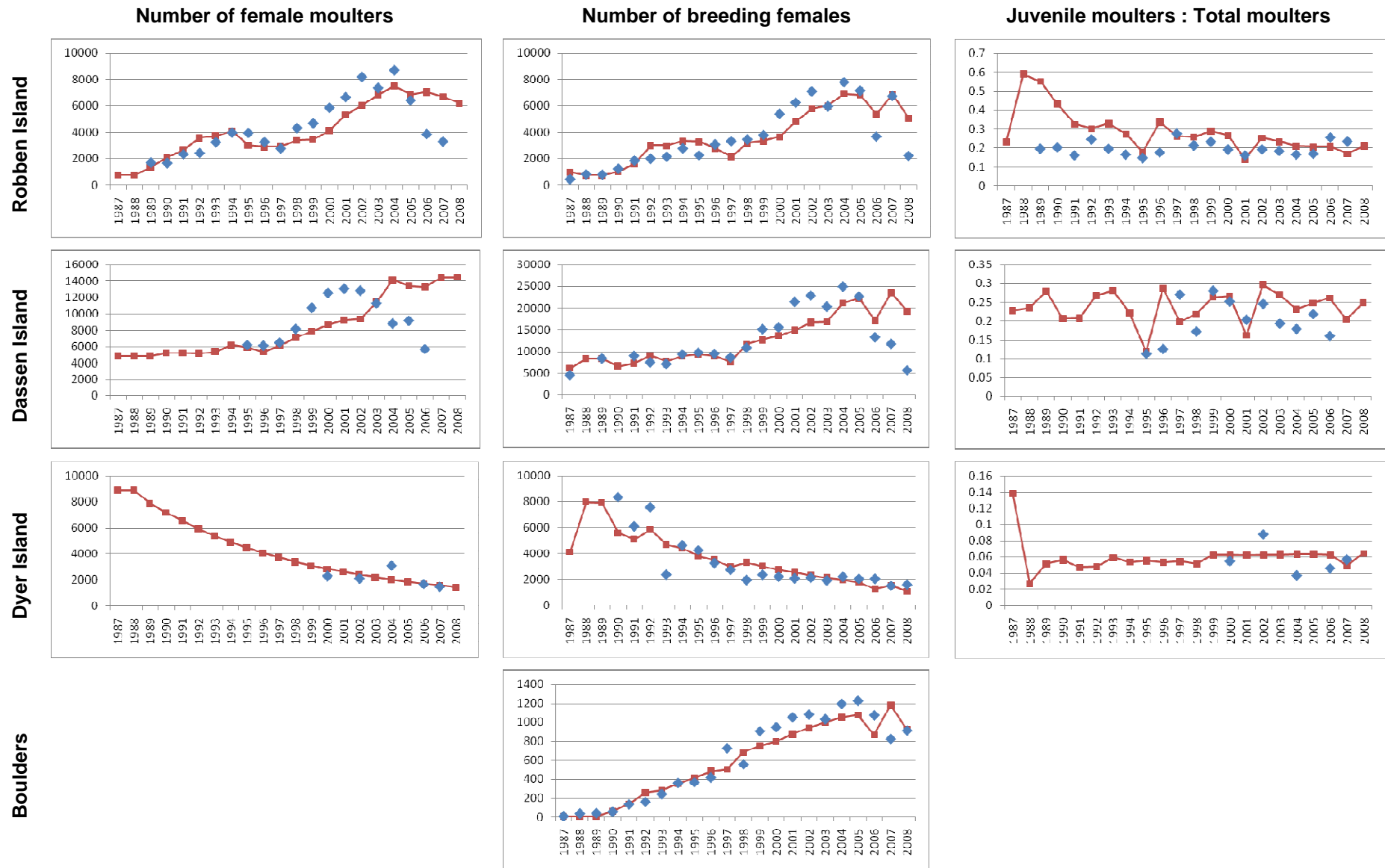


Fig. 6. Illustrative results from the final current Stage III model with adult survival dependent on pelagic abundance based on the relationship determined from the Stage II model (See Fig. 5).

## APPENDIX A

### POPULATION MODEL

#### An age-structured dynamic model

The basic dynamic model used to represent the population dynamics of the female penguins when assuming no movement between sub-areas is as follows:

$$N_{y+1,a+1}^s = N_y^{breed,s} q_f \cdot \omega_1 \cdot (S_y^{j,s})^{3/4} \quad a=0 \quad (1)$$

$$N_{y+1,a+1}^s = N_{y,a}^s \cdot (S_y^{j,s})^{1/4} S_y^{3/4} \quad a=1 \quad (2)$$

$$N_{y+1,a+1}^s = N_{y,a}^s \cdot S_y \quad 2 \leq a < 4 \quad (3)$$

$$N_{y+1,m}^s = (N_{y,m}^s + N_{y,m-1}^s) \cdot S_y \quad m=5 \quad (4)$$

with

$$N_y^s = \sum_{a=0}^m N_{y,a}^s \quad (5)$$

$$N_y^{breed,s} = \omega_2 \cdot \sum_{a=2}^m v_{y,a} \cdot N_{y,a}^s \quad (6)$$

$$S_y^{j,s} = S_j^* \left( 1 - \frac{\sum_{a=2}^m N_{y,a}^s}{K_s^*} \right) \quad (7)$$

$$\omega_1 = f_1(B_y) \cdot H_{\max} \quad (8)$$

$$\omega_2 = f_2(B_{y-1}) \cdot 1 \quad (9)$$

where:

- $N_{y,a}^s$  is the number of female penguins of age  $a$  on 1 April in sub-area  $s$  and year  $y$ ;
- $N_y^{breed,s}$  is the number of breeding penguins in area  $s$  and year  $y$ ;
- $q_f$  is the fraction of chicks that are female;
- $S_y$  is the post-first-year survival rate in year  $y$ ;
- $S_y^{j,s}$  is the annual post-fledging survival rate of juvenile penguins in year  $y$  and sub-area  $s$ ;
- $m$  is the largest age considered (the “plus group”, set at 5 years);
- $v_{y,a}$  is the proportion of female penguins of age  $a$  that mature and commence breeding in year  $y$ , determined as a function of a component of pelagic fish abundance  $B$  in year  $y-1$ ;
- $f_1(B_y)$  is a breeding success factor (multiplier for  $H_{max}$ ) to compute the average fledging success, and which is a function of a component of pelagic fish abundance  $B$  in year  $y$ ;
- $H_{max}$  is the maximum observed breeding success (chicks fledged per female per yr);
- $f_2(B_{y-1})$  is a factor determining the proportion that breed in year  $y$  (with the maximum possible breeding proportion set at 1) which is a function of a component of pelagic fish abundance  $B$  in year  $y-1$ ; and
- $K_s^*$  is a carrying capacity-related term for adult penguins in sub-area  $s$ , used to introduce density dependence into the penguin dynamics through the dependence of  $S_y^{j,s}$  on the total abundance of all penguins aged 2 and older.

### Breeders and moulters

Equation (6) above provides a model value for the numbers of breeding females in each of the sub-areas. The numbers of female juvenile and adult moulters present on the 1 December each year  $y$  are given respectively by:

$$N_{y,Juv\_moult}^s = N_{y,1}^s \cdot (S_y^{j,s})^{1/4} S_y^{5/12} \quad (10)$$

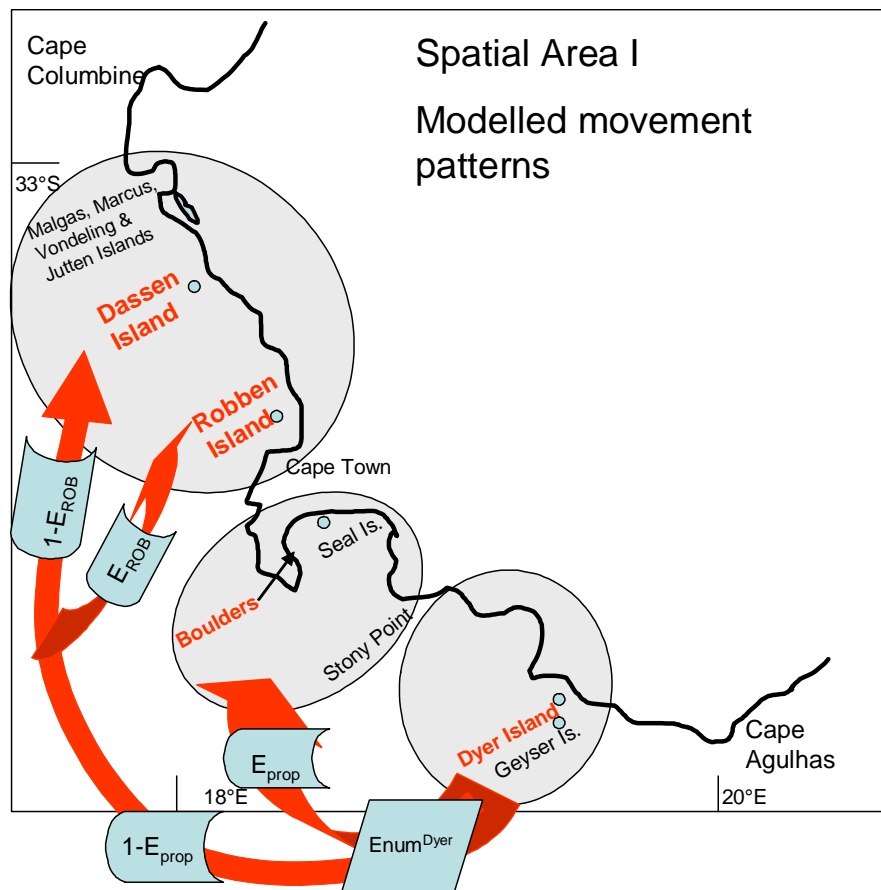
$$N_{y,Ad\_moult}^s = \sum_{a=2}^m N_{y,a}^s \cdot S_y^{2/3} \quad (11)$$

and hence the juvenile moulters as a proportion of total moulters is computed as:

$$P_{y,Juv:Tot} = N_{y,Juv\_moult}^s / (N_{y,Juv\_moult}^s + N_{y,Ad\_moult}^s) \quad (12)$$

### Adding immigration and emigration effects

The model allows emigration from Dyer Island to Boulders and to Robben and Dassen Islands. The initial movement model is a simple one and could be improved in future. It assumes that a fixed proportion (estimated within the model) of one-year-olds from Dyer move west each year. Of these birds moving west, a second proportion (again estimated within the model) are assumed to immigrate to Boulders. The trend in numbers at Boulders can be explained only by immigration and the model estimates the proportion of the westward emigrants that are needed to match the observed trends at Boulders. The remainder of the westward emigrants move to Robben and Dassen Islands, with the relative proportion moving to each island again estimated within the model.



Mathematically, for Dyer Island, Equation (3) for the case  $a=2$  is thus modified as follows:

$$N_{y+1,a+1}^s = N_{y,a}^s \cdot S_y \cdot (1 - E_y) \quad (13)$$



where  $E_y$  is the proportion of first-time breeders (assumed to be between the ages of 1 and 2 years) that emigrate from Dyer Island. The emigration proportion is estimated (as a constant year-independent value) within the model. The actual number of birds emigrating each year  $Enum_y^{Dyer}$  is thus:

$$Enum_y^{Dyer} = E_y \cdot S_y \cdot N_{y,2}^s \quad (14)$$

The proportion of the first time breeders that immigrate to Boulders (Bld) rather than to Robben and Dassen islands is determined by parameter  $E_{prop}$ , estimated within the model. Similarly, the proportion of these (i.e. the balance that move to Robben or Dassen) is determined by the parameter  $E_{ROB}$ , also estimated within the model. It follows that Equation (3) (for the case  $a=2$ ) must be modified as follows for the Robben, Dassen and Boulders colonies respectively:

$$N_{y+1,3}^{Rob} = N_{y,2}^{Rob} S_y + (1 - E_{prop}) \cdot (E_{ROB}) \cdot Enum_y^{Dyer} \quad (15)$$

$$N_{y+1,3}^{Das} = N_{y,2}^{Das} S_y + (1 - E_{prop}) \cdot (1 - E_{ROB}) Enum_y^{Dyer} \quad (16)$$

$$N_{y+1,3}^{Bld} = N_{y,2}^{Bld} S_y + E_{prop} \cdot Enum_y^{Dyer} \quad (17)$$

### Taking account of major oil spills

The Apollo Sea oil spill in 1994 and Treasure oil spill in 2000 resulted in the death of approximately 5000 and 2000 breeding adults, mostly from Robben and Dassen islands (Underhill *et al.* 1999, 2006, Crawford *et al.* 2000). As this is an important additional source of mortality, in the model it is assumed that an additional 2500 and 1000 breeding females from Dassen and Robben died in these years, with the number assumed dead from each of these two colonies determined by a fixed input parameter ( $oil\_prop$ ; set = 0.5 in illustrative example) and from each age class computed on the assumption of proportionality to the abundance of that age class.

Equations (3) (and Equation (4) similarly) for Robben and Dassen is thus modified as follows:

$$N_{y+1,a+1}^s = N_{y,a}^s \cdot S_y - p_a M_y^{oil} \quad (18)$$

where

$$P_a = \frac{N_{y,a}^s}{\sum_{a'=2}^m N_{y,a'}^s}$$

and  $M_y^{oil}$  is an additional mortality term set as the observed number of penguins age 2+ that died in year  $y$  as a result of oiling.

In addition, it is likely that most if not all chicks fledged in those years will have died (Crawford, pers. commn) and hence it is assumed that only a small proportion (set at 50% here) of age 1 birds survived in those years.

### Annual variation in adult survival rate

The simplest version of the model assumes adult survival  $S$  is constant over time. Allowing adult survival rate to vary freely from year-to-year as an annual estimable parameter (i.e.  $S \rightarrow S_y$ ) would lead to an over-parameterised model, but process error of this kind can be incorporated in the model by treating it as a random effect. To maintain the  $S_y \leq 1$  constraint, it is easiest to transform to an annual adult natural mortality rate ( $M_y$ ):

$$M_y = -\ln S_y \quad (19)$$

where  $M_y = M e^{\eta_y}$   $\eta_y \sim N(0, (\sigma_\eta)^2)$  (20)

and the  $M$  and  $\eta_y$ 's become estimable parameters.

In a strict frequentist paradigm, obtaining maximum likelihood estimates requires integration over these random effects. The simpler approach here may be thought of as providing estimates for the Bayesian posterior modes, given uniform priors for all the estimable parameters except the  $\eta_y$ 's for which normally distributed priors of mean zero and variance  $\sigma_\eta^2$  are assumed (see Equation (20)).

In the model, a common baseline natural mortality rate is estimated. It is assumed that mortality at Dyer and Boulders is fixed at this value, while the mortality at the Robben and Dassen colonies may vary annually. Initially the same estimated residuals were used for the two islands, but this method failed to produce satisfactory fits to the moult and breeder data for Robben and Dassen

simultaneously. For this reason, an attempt was made to estimate separate sets of  $\eta_y$ 's for the two islands.

With a view to projecting penguin population numbers into the future, we explored the nature and strength of a functional relationship between the survival residuals and a pelagic abundance index. Examining correlations between the estimated residuals and various pelagic indices, a strong linear relationship was seen for Robben Island (the best was with the sardine and anchovy combined recruitment index,  $r^2=0.44$ ), but no statistically significant correlation was found for Dassen Island. The lack of a common relationship for both colonies was very evident when trying to refit the model using the functional relationships for the survival instead of the more flexible, residuals approach (Fig. 6).

### Starting values and equilibrium assumptions

Given a value for the starting (1987) number of adult moulters (ages 2+) in an area  $N_{0,Ad\_moult}^s$ , the total initial population size can be computed as well as the initial number of breeding females. Under the assumption that the population of each colony is at equilibrium in the year the model commences, and assuming further that there is zero initial immigration/emigration, as well as that  $\omega_1 = H(1989)$  (i.e. fledging success observed in 1989) and  $\omega_2 = \omega_2(1989)$  (i.e. proportion of mature birds expected to breed that year), simplifies the problem of solving for the starting number of breeding females  $N_0^{breed,s}$  through solving the balance equation:

$$N_0^{breed,s} = N_{0,Ad\_moult}^s \cdot \left( q_f \omega_1 (S_0^{j,s})^{\frac{3}{4}} \right)^{-1} \cdot \left( (S_0^{j,s})^{\frac{1}{4}} \cdot S_0^{\frac{3}{4}} \cdot \left( 1 + \sum_{a=3}^{m-1} S_0^{a-2} + \frac{S_0^{m-2}}{(1-S_0)} \right) \right)^{-1} \quad (21)$$

with

$$S_0^{j,s} = S_j^* \left( 1 - \frac{N_{0,Ad\_moult}^s}{K_s^*} \right) \quad (22)$$

Given  $N_0^{breed,s}$ , the initial numbers at each age  $a$  can then be computed as:

$$N_{0,a}^s = \begin{cases} N_0^{breed,s} \cdot q_f \cdot \omega_1 \cdot (S_0^{j,s})^{\frac{3}{4}} & a = 1 \\ N_0^{breed,s} \cdot q_f \cdot \omega_1 \cdot (S_0^{j,s}) S_0^{\frac{3}{4}} & a = 2 \\ N_0^{breed,s} \cdot q_f \cdot \omega_1 \cdot (S_0^{j,s}) S_0^{\frac{3}{4}} S_0^{a-2} & 3 \leq a < m \\ N_0^{breed,s} \cdot q_f \cdot \omega_1 \cdot (S_0^{j,s}) S_0^{\frac{3}{4}} S_0^{m-2} / (1 - S_0) & a = m \end{cases} \quad (23)$$

### Fitting procedure

The estimable parameters are estimated by minimising the total negative log likelihood. The model is fitted to moult count data (available for Robben, Dassen and Dyer Islands), breeding numbers data (available for Robben, Dassen, Dyer and Boulders) and to data on the proportion of juvenile to total moulters (available for Robben, Dassen and Dyer). Contributions by each of these to the negative of the log-likelihood ( $-\ln L$ ) are as follows.

### Moult counts

The moult counts are assumed to be log-normally distributed indices of the total numbers of adult female penguins such that:

$$I_y^{moult,s} = \hat{I}_y^s e^{\varepsilon_y^s} \quad \text{or} \quad \varepsilon_y^s = \ln(I_y^{moult,s}) - \ln(\hat{I}_y^s) \quad (24)$$

where  $I_y^{moult,s}$  is the (observed) moult count (expressed in terms of females only) for sub-area  $s$  and year  $y$ ;

$\hat{I}_y^s = q^s N_{y,Ad\_moult}^s$  is the corresponding model estimated value, where  $\hat{N}_{y,Ad\_moult}^s$  is the model-estimate of the number of female moulters, as given by Equation (11);

$q^s$  is the constant of proportionality for moulters corresponding to sector  $s$ ; and

$$\varepsilon_y^s \quad \text{from} \quad N\left(0, (\sigma_y^s)^2\right).$$

The contribution of the moult count data to the negative of the log likelihood function (after removal of constants) is given as shown below with contributions added over sub-areas  $s$ :

$$-\ln L^{moult} = \sum_s \left[ n_s \ln(\hat{\sigma}_s^{moult}) \right] + \frac{n_s}{2} \quad (25)$$

with

$$\hat{\sigma}_s^{moult} = \sqrt{\left[ \sum_{y=1}^{n_s} (\ln I_y^{moult,s} - \ln \hat{I}_y^s)^2 \right] / n_s} \quad (26)$$

where

$n_s$  is the number of years for which there are moult count data in sub-area  $s$ .

The proportionality constant  $q^s$  for moult numbers in sub-area  $s$  is estimated by its maximum likelihood value:

$$\ln \hat{q}^s = \frac{1}{n_s} \sum_y (\ln I_y^s - \ln \hat{N}_{y,Ad\_moult}^s) \quad (27)$$

### **Breeding pairs**

The breeding pairs data are assumed to be log-normally distributed indices of the total numbers of adult female penguins that breed each year, such that:

$$I_y^{breed,s} = \hat{I}_y^s e^{\varepsilon_y^s} \quad \text{or} \quad \varepsilon_y^s = \ln(I_y^{breed,s}) - \ln(\hat{I}_y^{br,s}) \quad (28)$$

where  $I_y^{breed,s}$  is the (observed) breeding pairs count for sub-area  $s$  and year  $y$ ;

$\hat{I}_y^{br,s} = q_{breed}^s N_y^{breed,s}$  is the corresponding model estimated value, where  $\hat{N}_y^{breed,s}$  is the model-estimate of the number of breeders, as given by Equation (6);

$q_{breed}^s$  is the constant of proportionality for breeders corresponding to sector  $s$ ; and

$\varepsilon_y^s$  from  $N(0, (\sigma_y^s)^2)$ .

The contribution of the breeding count data to the negative of the log likelihood function (after removal of constants) for all sub-areas  $s$  is thus as follows:

$$-\ln L^{breed} = \sum_s \left[ n_{b,s} \ln(\hat{\sigma}_s^{breed}) \right] + \frac{n_{b,s}}{2} \quad (29)$$

with

$$\hat{\sigma}_s^{breed} = \sqrt{\left[ \sum_{y=1}^{n_s} \left( \ln I_y^{breed,s} - \ln \hat{I}_y^{br,s} \right)^2 \right] / n_{b,s}} \quad (30)$$

where

$n_{b,s}$  is the number of years for which there are breeding count data in sub-area  $s$ .

The proportionality constant  $q_{breed}^s$  for Robben Island is assumed equal to 0.9 (to reflect minor undercounting) and the  $q_{breed}^s$  values for the other colonies are set equal to the Robben Island value.

### ***Proportion of juveniles:***

The log-likelihood contribution from the juvenile proportional abundance data is given by:

$$-\ln L^{prop} = \sum_s \left[ n_{p,s} \ln(\hat{\sigma}_s^{prop}) \right] + \frac{n_{p,s}}{2} \quad (31)$$

with:

$$\hat{\sigma}_s^{prop} = \sqrt{\left[ \sum_{y=1}^{n_{p,s}} \left( I_y^{Prop,s} - \hat{p}_{y,Juv:Tot}^s \right)^2 \right] / n_{p,s}} \quad (32)$$

where

$I_y^{Prop,s}$  is the (observed) juvenile proportional abundance data for sub-area  $s$  and year  $y$ ;

$n_{p,s}$  is the number of years for which there are proportional abundance data in sub-area  $s$ ; and

$\hat{p}_{y,Juv:Tot}^s$  is the model-estimate of the proportion of juvenile moulters to the total number of moulters, as given by Equation (12).

### ***Adult mortality residuals:***

When estimating residuals, the following penalty term is added to the negative log likelihood:

$$-\ln L^{resid} = \sum_s \sum_y (\eta_y^s)^2 / (2\sigma_\eta^2) \quad (33)$$

where

$\eta_y^s$  is the adult survival residual for year  $y$  in sector  $s$ , which is estimated for years  $y_1$  to  $y_2$  (set here as 1988 to 2008); and

$\sigma_\eta$  is the standard deviation of the log-residuals, which currently is input with a value of 0.6. (The sensitivity of results to alternatives for this choice will subsequently be checked).

### ***Parameter estimates and constraints***

Parameters estimated in the model were constrained in the following ranges for reasons of biological realism:

$$\begin{aligned} 0 < M &\leq 1.0 \\ 0 &\leq E \leq 1.0 \\ 0 &\leq E_{prop} \leq 1.0 \\ 0 &\leq E_{ROB} \leq 1.0 \\ 0 &< N_{0,Ad\_moult}^{Rob} \\ 0 &< N_{0,Ad\_moult}^{Das} \\ 0 &< N_{0,Ad\_moult}^{Dy} \end{aligned}$$

Apart from the 7 parameters above, an additional 42  $\eta_y^s$  parameters are estimated during the first stage when seeking a correlation between survival and pelagic abundance.

## APPENDIX B

### ASSUMED RELATIONSHIPS WITH FOOD AVAILABILITY

#### Breeding proportion

It has been suggested (R. Crawford, pers. commn) that the proportion of birds that choose to breed in a particular year may vary depending on food availability, as reflected by the  $f_2$  functional dependency of Equation (9). A number of different formulations are possible regarding the change in the proportion of mature birds that choose to breed (in April) each year  $y$  as a function of relative prey biomass  $B^r$  the previous year  $y-1$ , where prey biomass  $B^r$  is computed as relative to the maximum value observed over a series (Fig. B1). The simplest assumption is that 100% of mature birds breed each year. At the other extreme, the proportion breeding may be assumed a simple linear function of prey biomass, but this was not considered realistic by the task group. The example here assumes 100% of birds breed if relative prey biomass is above the median  $B^r$  value, but that the proportion breeding decreases linearly for lower values of combined anchovy and sardine biomass with a one year time lag.

#### Age of maturation

Crawford (pers. commn.) has also suggested that the proportion of birds of different ages that mature each year  $y$  may be a function of relative prey biomass  $B^r_{y-1}$  the previous year  $y-1$ , where prey biomass is computed as relative to the maximum value observed over a series (e.g. Fig. B2). The simplest model assumption is that the base-case input values do not change over time. The current model assumes that all birds of age 3 and older are mature.

#### Chick fledging success

One of the most important functional relationships that needs to be determined before the pelagic OMP testing process can take account of the relationship between the breeding success of African penguins and the abundance of pelagic fish concerns the way in which the average numbers of chicks fledged each year relates to pelagic fish biomass, i.e. the  $f_1$  functional dependence in Equation (8). The model can currently be run either by using the observed values for Robben and Dassen Island, or by fitting a functional relationship within the likelihood maximisation. The latter should result in approximately the same level of variability as evident from the Robben island data and the maximum average number of chicks fledged is assumed equal to the maximum observed values.



There are a number of ways in which this relationship could be modelled. Following Plagányi and Butterworth (2006), a breeding success factor  $f_1(B_y)$  is thus formulated as a function of the available biomass of prey and acts as a multiplier to the maximum observed chick fledging rate (Fig. B3). To reduce the number of parameters in the model, the breeding success factor is scaled such that it is 1 when the pelagic biomass is at the maximum observed value. A useful functional form to use is that classically referred to as a Beverton-Holt stock-recruitment relationship, modified here to represent fledging success as a function of pelagic biomass,  $B_y$ :

$$f(B_y) = \frac{\alpha \cdot B_y}{\beta + B_y} \quad (33)$$

where  $\alpha$  and  $\beta$  are parameters with  $\beta = (\alpha - 1)$  so that  $f(1) = 1$ .

When adding a term to allow for fluctuations about this relationship, Equation (33) becomes:

$$f(B_y) = \frac{\alpha \cdot B_y}{(\alpha - 1) + B_y} e^{(\zeta_y - \sigma_{BR}^2/2)} \quad (34)$$

where

$\zeta_y$  reflects fluctuation about the expected curve in year  $y$ , which is assumed to be normally distributed with standard deviation  $\sigma_{BR}$ . (These fluctuations are not employed in the current model.)

In order to work with estimable parameters that are more meaningful biologically, Equation (33) is re-parameterised in terms of the maximum pelagic biomass,  $B_{MAX}^r$ , and the “steepness”  $h$  of the relationship, where “steepness” is the fraction of maximum fledging success that results when  $B_y^r$  drops to 20% of its maximum level, from which it follows that:

$$h = \frac{\alpha}{5\alpha - 4} \quad (35)$$

By ignoring the random variation term and choosing a single parameter value  $h$ , the fledging success relationship can thus be set. The parameter  $h$  may be thought of as controlling the level of prey biomass below which breeding success is negatively impacted (Fig. B3).

When fitting the chick fledging success functional relationship, the contribution to the negative of the log likelihood function (after removal of constants) is computed by summing over the Robben and Dassen Island contributions as follows:

$$-\ln L^{fledge} = \sum_s \left[ n_{f,s} \ln(\hat{\sigma}_s^{fledge}) \right] + \frac{n_{f,s}}{2} \quad (36)$$

with

$$\hat{\sigma}_s^{fledge} = \sqrt{\frac{\sum_{y=1}^{n_{f,s}} (\ln I_y^{fledge,s} - \ln \hat{I}_y^{fledge,s})^2}{n_{f,s}}} \quad (37)$$

where  $I_y^{fledge,s}$  is the (observed) mean fledging success for sub-area  $s$  and year  $y$ ;

$\hat{I}_y^{fledge,s}$  is the corresponding model-estimated value, as given by Equation (33) (or (34)); and

$n_{f,s}$  is the number of years for which there are fledging success data in sub-area  $s$ .

For illustrative purposes, in the results presented here, chick fledging success is calculated using the functional form of Equation (33), with separate parameters estimated for Dassen Island and Robben Island. The maximum values are set to 0.97 for Robben and 1.37 for Dassen. Fledging success is assumed equal to 1.0 for Dyer and Boulders.

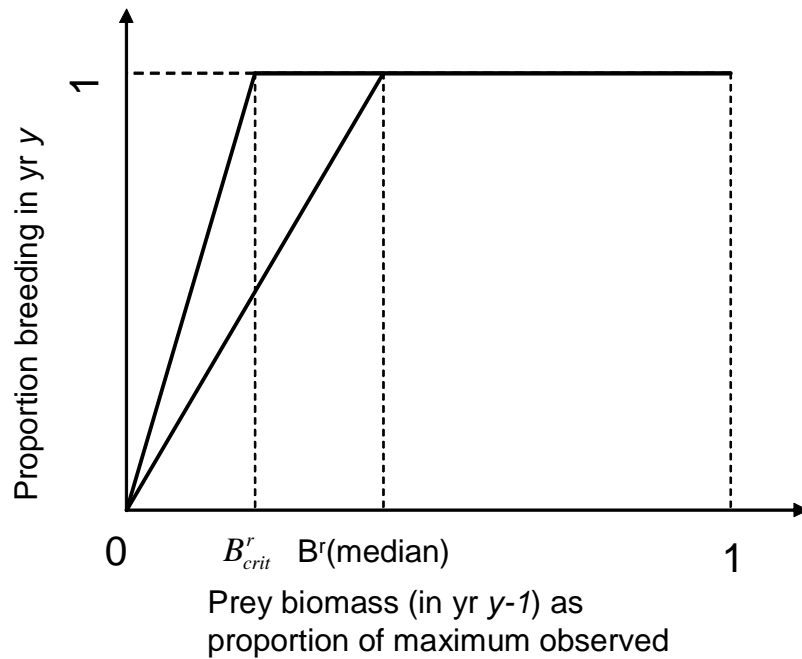


Fig. B1. Schematic showing different possible assumptions regarding the change in the proportion of mature birds that choose to breed (in April) each year  $y$  as a function of relative prey biomass  $B^r$  the previous year  $y-1$ , where prey biomass is computed as relative to the maximum value observed over a series. The simplest assumption is that 100% of mature birds breed each year. Other options would be to assume 100% of birds breed if relative prey biomass is above the median  $B^r$  value or an estimated  $B^r_{crit}$  value, but that the proportion breeding decreases linearly for lower values of  $B^r$ .

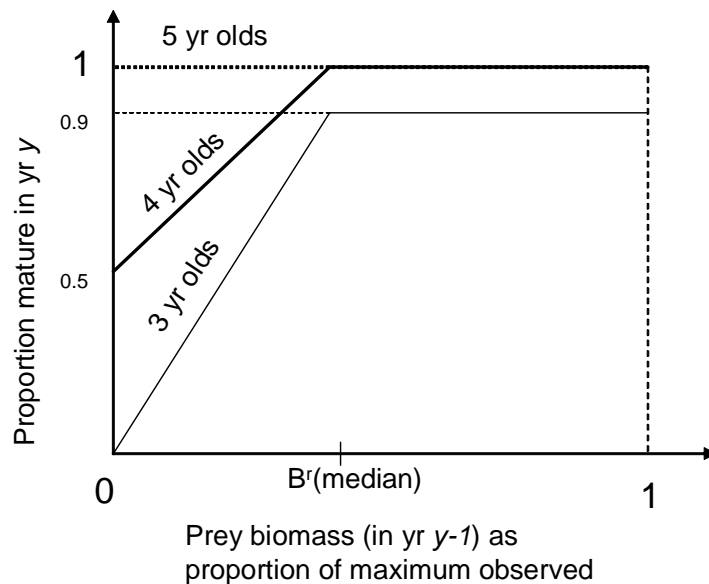


Fig. B2. Schematic showing initial input assumption regarding the change in the proportion of birds of different ages that mature each year  $y$  as a function of relative prey biomass  $B^r$  the previous year  $y-1$ , where prey biomass is computed as relative to the maximum value observed over a series. The simplest assumption is that these proportions do not change over time. An option being tried in the model involves assuming that the proportion maturing at ages 3 and 4 remains fixed at 0.9 and 1.0 respectively if relative prey biomass is above the median  $B^r$  value, but then decreases linearly with lower values of  $B^r$  as shown.

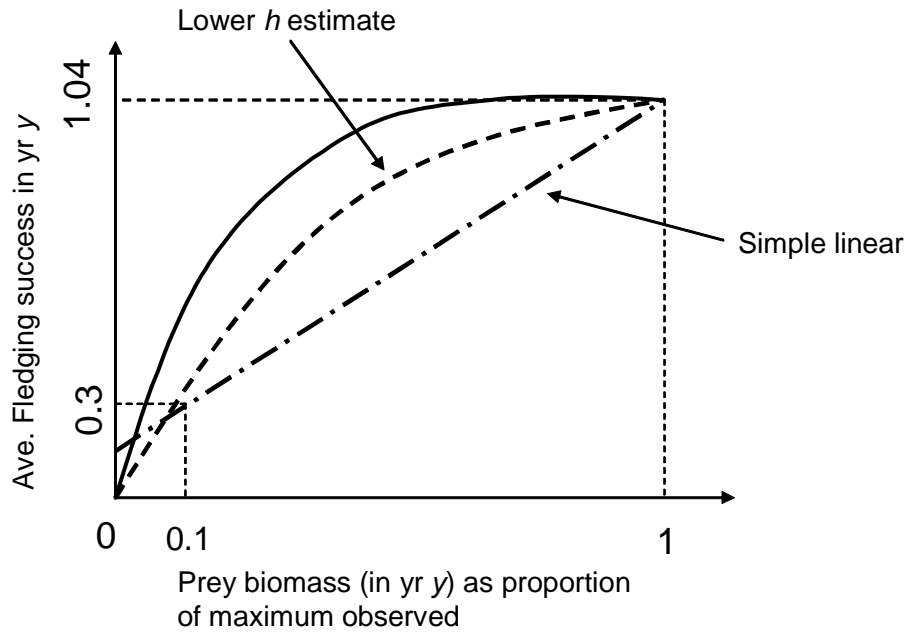


Fig. B3. Schematic showing different possible assumptions regarding the change in the average fledging success each year  $y$  as a function of relative prey biomass  $B^r$  in the same year  $y$ , i.e. the  $f_l$  functional dependence in Equation (8). Prey biomass is computed as relative to the maximum value observed over a series. The shape of the curve is determined by a single “steepness” parameter  $h$  with different values leading to either a near-linear decrease in fledging success as prey abundance decreases versus a scenario in which fledging success drops below the maximum observed level only at relatively low levels of prey abundance.