

## **DETECTING TRENDS IN ALBATROSS POPULATIONS**

**Miguel Pascual**

Centro Nacional Patagónico, CONICET.

Blvd. Brown S/N

(9120) Puerto Madryn, Chubut, ARGENTINA

Phone: 54.2965.451024 Fax: 54.2965.451543

E-mail: pascual@cenpat.edu.ar

**Ann E. Edwards**

Department of Zoology, Box 351800

University of Washington

Seattle, WA 98195, USA

E-mail: aee@u.washington.edu

**A Report Prepared for the New Zealand Hoki Fishery Management Co Ltd**

**November 1999**

# Table of contents

<b>EXECUTIVE SUMMARY</b> .....	<b>2</b>
<b>INTRODUCTION</b> .....	<b>4</b>
<b>A PRELIMINARY ANALYSIS</b> .....	<b>5</b>
<b>OVERVIEW OF APPROACH</b> .....	<b>7</b>
<b>METHODS</b> .....	<b>9</b>
ALBATROSS BIOLOGY .....	9
POPULATION MODEL.....	14
TREND ANALYSIS.....	15
SCENARIOS SIMULATED IN THIS REPORT.....	17
<b>RESULTS</b> .....	<b>18</b>
MODEL VALIDATION.....	18
POWER ANALYSIS FOR TRENDS IN COUNTS OF BREEDING PAIRS .....	21
POWER ANALYSIS WITH RESAMPLING METHODS .....	25
OTHER INDICES OF POPULATION SIZE .....	26
<b>DISCUSSION</b> .....	<b>28</b>
<b>REFERENCES</b> .....	<b>31</b>
<b>APPENDIX 1: ALBATROSS DATA</b> .....	<b>33</b>
<b>APPENDIX 2: MODEL</b> .....	<b>40</b>
<b>APPENDIX 3: GUIDE TO SPREADSHEET USE</b> .....	<b>43</b>
<b>APPENDIX 4: PARAMETERS DEFINING THE SCENARIOS USED IN THE SIMULATIONS.</b> .....	<b>46</b>

## Executive Summary

- Breeding populations of Wandering Albatrosses have been declining. Populations on Adams and Antipodes Islands, New Zealand, constitute a large proportion of the total number of great albatrosses (*Diomedea* spp.) found in the southern oceans, yet systematic monitoring of the New Zealand populations began only in the 1990s. Estimating population level trends in long-lived, biennially breeding species such as albatrosses requires many years of monitoring data. However, there is a significant short-term management need for information about population trends, and the demographic factors that most strongly influence those trends. In this project, we developed and explored a series of approaches intended to maximize the likelihood of detecting population level changes in biennially breeding albatrosses.
- We developed and implemented a stochastic population dynamics model of albatross species. The model is general enough to accommodate a broad range of life history characteristics, and sampling protocols.
- We applied the model to the analysis of statistical power that different statistical methods provided to detect trends in population numbers. We derived some general conclusions about how much data are needed to detect trends of different sizes and explored the effects of sampling error on our ability to detect trends. We also scrutinized the performance of alternative statistical methods.
- Of the three statistical methods used, a maximum likelihood estimation approach failed completely to detect trends in numbers, while a simple linear regression of  $\ln(\text{numbers})$  on time provided higher power.
- In general, 10 years of data provided very low power to detect trends of 1% a year. The life-history driven, inter-annual variability in the size of the breeding population, together with sampling errors and inherent individual and environmental variation, create the need for longer time series for attaining significant power. At least 15 years of data will be required to detect declining trends of 1% per year in populations of biennially breeding albatross species when sampling errors are insignificant, but power deteriorates quickly as sampling errors increase.
- We implemented a randomization approach based on resampling blocks of data, which worked significantly better than the simple regression approach. This method, however, also is strongly affected by the addition of sampling error.

- We also explored the value of counting old juveniles at the colonies as an index of abundance. Such an index is not expected to display the inherent annual fluctuations characteristic of counts of breeding adults. The value of juvenile counts as an index of abundance will depend on how consistently they return to the colony.
- We provide programs, documentation, and guidelines for users interested in applying these approaches to specific cases.

## Introduction

Subantarctic ecosystems have been affected by over a century of sealing and whaling in the circumpolar reaches of the Southern Oceans (Jouventin and Weimerskirch 1991). Most populations of seabirds inhabiting the remote islands in this immense region remained virtually untouched. Recently, the situation has changed for some species, as the nature and extent of human activities has changed. Several demographic studies have indicated that some southern seabird populations are declining (reviewed in Jouventin and Weimerskirch 1991, and Bergin 1997). The factors that are believed responsible for these declines include changes in climatic and oceanographic conditions, introduction of predators, and interaction with fisheries, both by competition for food and through by-catch mortality in fishing operations.

Regardless of the factors affecting southern seabird species, monitoring a population is logistically difficult, and documenting population trends requires expensive, long-term sampling protocols. Mathematical modeling, and simulation of population behavior given various demographic parameters and sampling protocols, offer the advantage of evaluating the probabilities for detecting population level trends before launching extensive monitoring programs. Mathematical simulations also can provide crucial guidance in the development of robust sampling protocols.

The purpose of this project is to evaluate the probability of detecting trends in the abundance of albatross populations in New Zealand waters, with particular emphasis on Antipodean Albatross (*Diomedea antipodensis*) and Gibson's Albatross (*D. gibsoni*) populations on Antipodes Is. and Adams Is. A major management concern for these two species is the early detection of population level changes due to long-line associated by-catch mortality for vessels seeking tuna, ling, and tooth fish. The demographic information available for New Zealand albatross species is limited. Therefore, we draw heavily from data available for other *Diomedea* species, particularly populations of Wandering Albatross (*D. exulans*) on Bird Is., South Georgia, and on Possession Is., Crozet Islands. The tools developed for this project are designed for general use beyond New Zealand albatrosses, making them adaptable for analyzing population trends of other colonially breeding, marine organisms.

## A preliminary analysis

The probability of detecting a trend in population abundance will depend on: a) the intrinsic level and mode of variability in population numbers, b) the magnitude of the trend, c) the number of data points available, d) the magnitude and nature of sampling error, and e) the statistical power of the method used to detect the trend.

To illustrate how some of these factors affect detection of trends, we present the results of a simulation exercise based on Wandering Albatross data from South Georgia (Croxall et al., 1990). The data used here (Figure 1) consist of 20 annual counts of breeding pairs collected over a 28 year period. The data show a decrease of about 1.0% per year, a trend that is consistent with predictions from a model using detailed demographic data collected for these colonies (Croxall et al., 1990).

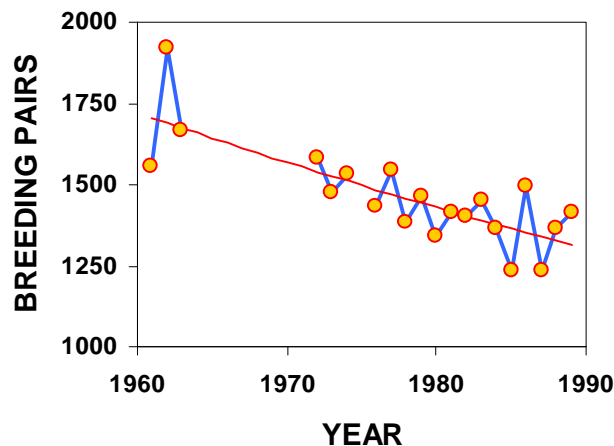


Figure 1. Time series of counts of breeding pairs of wandering albatross (*Diomedea exulans*) for South Georgia. The fitted regression line, a trend of  $-1\%$ , is shown. Source: Croxall et al 1990.

To estimate the probability of detecting trends we used a Monte Carlo simulation approach, analogous to a bootstrap (Efron and Tibshirani, 1993). We simulated population trajectories by resampling the annual breeding population data for South Georgia and evaluated the number of simulated data sets for which we were able to find a statistically significant decline. In order to do this, we first generated a deterministic decline of a given size (we tried 0.5%, 1%, 1.5% and 2%). In order to simulate variability in the counts consistent with that of the observations, we added to the deterministic counts residuals sampled with replacement from a fit of the original data (Figure 1). The “power” to detect a trend consists of the proportion of generated data sets for which the “null

hypothesis" (slope $\geq$ 0) was rejected with a one tail  $t$ -test ( $\alpha=0.05$ ).

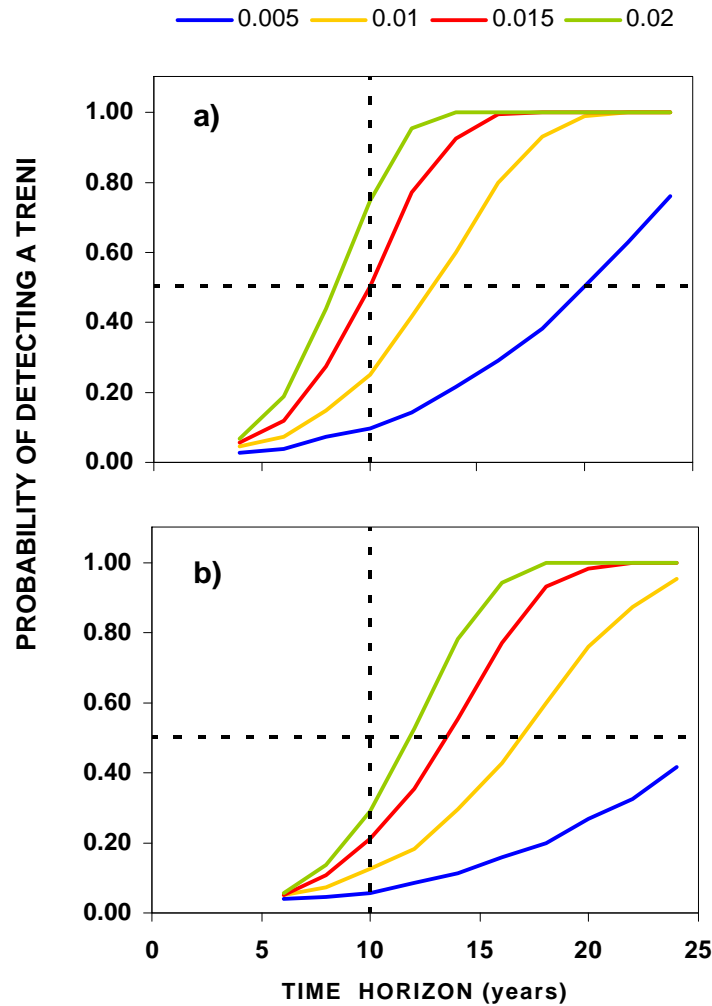


Figure 2. Power to detect trends of different size (different lines) over different time horizons when a) sampling every year, and b) sampling every other year.

Figure 2a shows the power to detect trends of different sizes (*i.e.*, different rates of change in population size) when the total number of years sampled is varied. As expected, for data collected over a given total number of years, the probability to detect a trend increases as the trend is more pronounced. In general, these simulations suggest that to attain a 80% probability of detecting a trend of 1% a year (such as that of the South Georgia population) 15 years of data are required. Ten years of data provide a low power to detect a trend of 1% (about 25%), but a much higher power is attained for the same amount of data if the trend is 2% per year (about 75%).

Figure 2b shows the same type of simulation, but instead of sampling every year, population censuses are simulated to occur every other year. These results show that the time horizon is very important in detecting a trend. For instance, 10 population censuses undertaken every other year over a 20 year period, produces as much power to detect a trend of 1% as 16 population censuses conducted annually (over 75%, which is significantly more power than that provided by 10 consecutive censuses).

These results show how trend size, number of years in the sampling period, and sampling periodicity, all affect the power to detect a trend. Resampling individual residuals provides a direct, uncomplicated approach and serves to illuminate the main factors that affect trend detection. It does not, however, provide a means for exploring the effects of certain other factors, such as the size and shape of variability as a function of particular biological characteristics or those of different sampling errors. As discussed later, this approach also fails to acknowledge the particular temporal patterns found in time series of breeding populations of biennial species. Therefore, we developed a more sophisticated simulation framework that allows us to account for such factors. We also explored alternative statistical analytical tools in search of approaches that provide maximum power to detect trends in albatross populations.

## **Overview of approach**

The primary tool we built for detecting population trends in albatrosses was a generalized stochastic model of albatross population dynamics. We used this model to simulate: a) population trajectories into the future, b) various “sampling” protocols, and c) assessment of trends with different statistical methods (Figure 3). This approach allows us to include and control population characteristics, the magnitude of the trend, the number of data points available, the amount and nature of sampling error, and the statistical method used to test for the trend. It also allows to consider increases in incidental mortality.



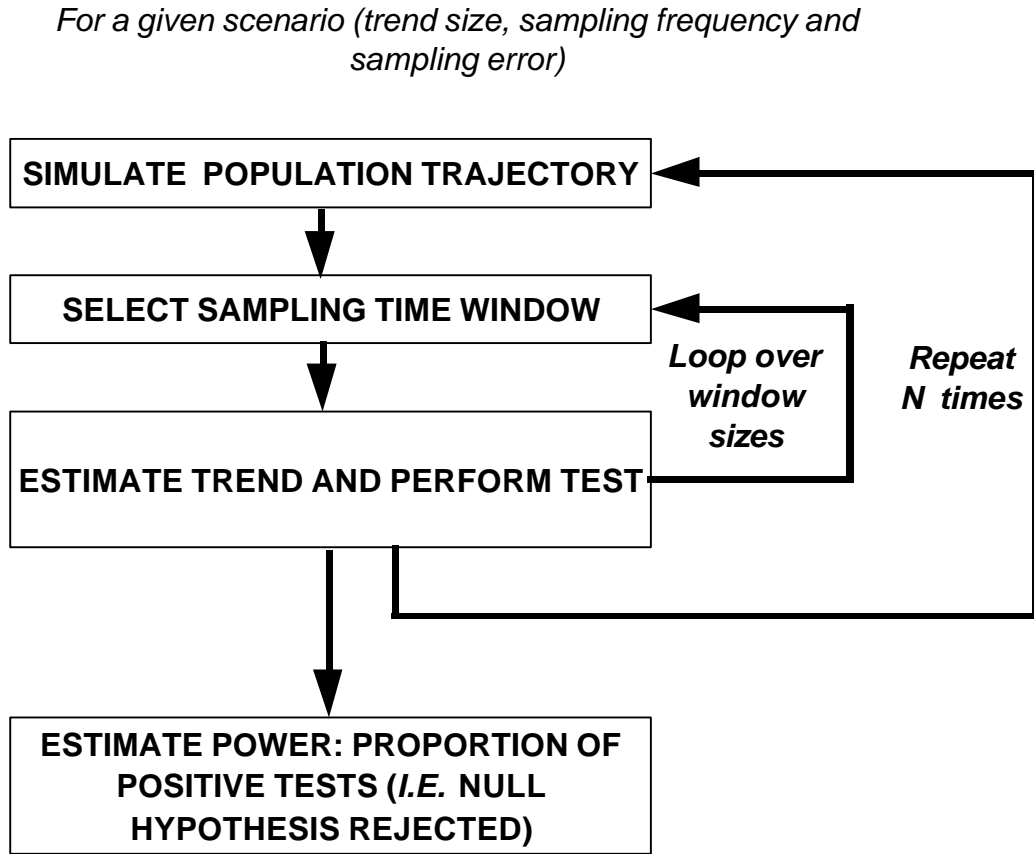


Figure 3. Flow chart of the simulation approach used to estimate the power to detect population trends.

Because the goal was to generate an analytical tool that would be applicable across populations, species and sampling designs, we intentionally built a general model, yet one that was small enough to provide meaningful results. The cost of making a flexible model is that it necessarily has a large number of parameters that define the biological characteristics, the sampling scheme and the statistical methods employed.

While the model was developed as an exploratory tool for trend detection, it also can be used to a) simulate population behaviors as a function of known parameters, b) explore dynamic scenarios that may have produced observed population behaviors, and c) assess the tolerance of a given population to incidental by-catch mortality. Those were not, however, the emphases of this project and the model is used in this report only as a tool to explore the problem of detecting population declines. Here we have reduced the potential complexity of both population behavior and sampling variability, and only report results from a subset of scenarios. We believe the scenarios reported cover the range of most likely possibilities for biennially breeding albatross

species.

## Methods

### *Albatross biology*

In 1995, albatross phylogeny was revised, increasing the number of recognized species from 15 to 24 (Robertson and Nunn, 1998). The Wandering Albatross (*Diomedea exulans*) was reclassified to give species status to four previously recognized subspecies: 1) the Wandering Albatross (*Diomedea exulans*), for which the majority of birds breed on South Georgia, and the Crozet Islands; 2) the Tristan Albatross (*D. dabbenena*), found only on the islands of Tristan da Cunha; 3) the Antipodean Albatross (*D. antipodensis*), found only on Antipodes and Campbell Islands, N.Z.; and 4) the Gibson's Albatross (*D. gibsoni*), found only on the Auckland Islands, N.Z. (Gales, 1998). Eggs and adult birds of *D. gibsoni* and *D. antipodensis* are smaller than *D. exulans* (Robertson, 1975) but the ecological or phylogenetic reasons for these differences are not known. Of the many populations of these four species, only the Wandering Albatross populations of South Georgia and the Crozet Islands have been well studied.

Albatrosses are long-lived species (40+ years), with low fecundity (a single egg per breeding attempt) and delayed reproductive maturity (8+ years). They nest almost exclusively on oceanic islands where they are colonial. They are strongly faithful to their mates and to their nest sites, returning to each for years or even decades. When not at the colony, they travel enormous distances, almost never touching land, in search of widely dispersed prey, primarily squid and fish.

Albatross natural history strongly dictates the nature of the demographic data that can be collected. When albatrosses are not breeding, or practicing breeding-related behaviors at the colony, they are found feeding at densities too low to monitor reliably. Therefore, estimates of population size, and adult and juvenile survival rates are derived from data collected at the breeding colonies. Albatrosses on colonies are easy to work with for two reasons. First, both adults and chicks are easy to approach and capture, so marking them with individually identifiable bands and rechecking the bands is not difficult. Second, they have a strong tendency to return to initiate breeding at the colony from where they fledged, and an even stronger tendency to return year after year to nest at or near the same site where they nested previously. Thus, researchers can monitor known individuals throughout their lifetimes, making estimates of age-specific survival possible.

However, not all albatrosses breed every year. This provides a significant handicap when attempting to estimate total population size based on annual counts of breeding pairs. Many albatross species breed biennially, meaning they skip one year after successfully fledging a chick (and sometimes two or three years) before attempting to breed again. Even in annually breeding albatross species, a minority of successful breeders will skip one or more years before attempting to breed again.

Quantifying the population dynamics of albatrosses is further handicapped by the inaccessibility of juveniles. Once birds fledge, they do not return to their natal colony until they are four or more years old. They do not begin to breed (thus forcing them to return to a known, markable location) until they are generally eight or more years old. Therefore, rates of recruitment to the breeding population can be measured on birds banded as chicks, but measuring age-specific survival rates of young juveniles is not possible.

Albatross chicks grow very slowly. A successful breeding season will last from nine to fourteen months, depending on the species, and to some extent, individual variation. The breeding season for *Diomedea* species lasts twelve to fourteen months. At colonies censused only once a year, the long time lag between egg laying and chick fledging can make it difficult to accurately measure annual breeding success. Albatrosses at high latitudes are synchronous breeders, meaning all current breeders lay, hatch or fledge their chicks at similar times each year. Strong reproductive synchrony is beneficial for collecting robust, long-term, demographic data, but requires that annual censuses of , say, number of breeding pairs with eggs, or number of chicks, be made at the same time each year.

The typical monitoring protocol for an albatross breeding colony has two main components: 1) colony-wide censuses of actively breeding pairs based on counts of nests with eggs or chicks, and 2) more intensive studies of banded birds and marked nests. Counts of active nests over a large geographic area, usually an entire island, are typically the first type of data collected on a colony. Counts of active nests, are available for all populations of *Diomedea* species of the Wandering Albatross group studied thusfar (Table A1-1; Fig. A1-1). In the absence of survival rates, breeding frequencies and the other types of demographic data described below, it is not possible to estimate accurately the size of the entire population of a particular breeding island. Therefore, “population estimates” at sites visited only briefly each year are limited to a count of active nests.

Historical surveys of remote oceanic breeding colonies were somewhat opportunistic, so

data were collected at different times of year. This makes it difficult to compare historical data between years. More recent, systematic monitoring programs aim to count the number of active nests at the same time each year. Generally, the goal is to count nests as soon as possible after all birds have laid. This minimizes the number of failed nests that would be missed in a census of active nests, and increases inter-annual reliability. For *Diomedea* populations on South Georgia, Crozet, Adams and Kerguelen Islands, counts of active nests take place from mid-January to early February, often on a specifically designated day. Counts on Antipodes in 1994 took place in February/March. Due to logistical constraints, nests with chicks were counted annually in April/May on Marion and Prince Edward Is, and in September/November on Gough Is. Censusing the population at this time of year consistently underestimates the annual abundance of breeders, since egg and chick mortality are not taken into consideration (Watkins, 1987).

Before censusing whole islands (e.g., Bird Is. in South Georgia, Adams, and Antipodes), which may be many square kilometers, each island is divided into blocks based on distinct geographical features. Then survey teams of about four people walk parallel transects spaced about 20-50 m apart, zig-zagging through tussock grasses and shrubs on medium-grade slopes, counting all active nests. Mature birds found on the ground but not on a nest, and unfledged chicks of the previous year also may be counted. For large islands, even this simple data collecting task is energy intensive and time consuming (requiring 48 person-days to cover the breeding grounds on Adams Is. in 1991). Inter-annual comparisons of non-breeding bird abundance are not reliable because the number of non-breeding birds encountered on the ground varies with weather and time of day, factors which are not controlled for during island-wide censuses (e.g., Walker et al., 1991, 1995).

The second standard way to collect data on albatross demographics is by following known, banded birds in order to measure age-specific survival rates, breeding frequency, individual breeding history, and age of first reproduction. When birds are banded as chicks, their exact age is known throughout their lives. When breeding birds are banded, their minimum age is known (minimum age of first reproduction plus the number of years since the bird was banded). On Possession Is., 3312 adult birds and 4750 chicks were banded between 1960-1994.

For large colonies visited once a year for short periods (e.g., Adams, Antipodes, Marion) a subset of the colony is designated the permanent intensive study area. In this area all nests are marked and all breeding birds and chicks are banded. For Bird Is. and Possession Is., where

sampling efforts are greater, all birds on the islands are banded. In subsequent years, every bird within the study area is checked for bands, and unbanded birds, including chicks, are banded.

Nest site fidelity allows subsampling within the colony to be representative. However, on Adams Is. between 1991 and 1993, only 3% of pairs used the same nest site and others moved their nest site an average of 30m (SD=46; n=74). On Bird Is., pairs moved an average distance of only 7m and 20% used the same nest site. On Possession Is., 23% used the same nest site (Walker et al., 1995).

Adult Survival refers to the proportion of banded adults alive in one year that are still alive the next year. However, because some birds don't return to the colony for up to 4 years, it is necessary to wait at least four years before a banded bird is considered dead. It is considered to have been dead since one year after it was seen last. Adult survival on Possession Is. was significantly lower from 1966-1976 than from 1977-1993. Female survival (93.5-94.6%) was lower than for males(94.7-96.8%) when overall survival was low, but when overall survival was high there was no difference in adult survival rate between the sexes (Weimerskirch et al., 1997). Between 1976 and 1986 on Bird Is., adult survival rates for males were consistently higher (mean = 2.3%) than for females (Croxall et al., 1990; Fig. A1-2, A1-3). On Adams Is., out of the seven banded birds that returned in 1993 without their mates, six were male, suggesting that females had lower survival rates (Walker et al., 1995).

Survival estimates for New Zealand species are higher than those for *D.exulans*. Estimated average survival for *D.antipodensis* on Antipodes Is. for the period 1994-96 was 99% (Walker et al., 1999) and 97% for *D.gibsoni* on Adams Is. for the period 1991-97 (Walker and Elliot 1999).

Juvenile Survival, as currently calculated in the literature (for Possession and Bird Is.), refers to the number of banded chicks surviving to age 5. These data provide an index of overall juvenile survival, which has increased significantly since 1966 (Fig. A1-4).

Age at First Breeding refers to the age at which a bird first has an egg in its nest (birds will practice courtship and breeding behaviors for a number of years before this). On both Bird and Possession Is. the age of first breeding declined from the 1960s to the late 1970s. Age of the youngest breeders fell from 10 to 7 years on Bird Is. and from 9 to 7 years on Possession Is. The modal age fell from 12 to 10 on Bird Is. and from 13 to 10 on Possession Is. (Croxall et al., 1990; Weimerskirch et al., 1997).

Recruitment Rate refers to the proportion of chicks fledged in a given year (i.e., cohort) that

produce an active nest. Recruitment rate, therefore, is a function of juvenile survival and age at first breeding combined. Published cohort recruitment data are available only for fledglings at one site (Bird Is.) for five different years (Croxall et al., 1990; Fig. A1-5).

Breeding Frequency refers to the annual frequency with which mature birds attempt breeding. As explained above, birds may wait as many as four years before returning to breed. Therefore, it is not possible to collect complete breeding frequency data in less than four years. Since obtaining breeding frequency data requires both intensive banding effort and a long-term study, so far the only published breeding frequency data are from Bird Is. (Croxall et al., 1990; Fig A1-6a, b).

Breeding Success refers to the proportion of total eggs laid that produce fledged chicks. Obtaining this exact number, however, requires intensive sampling effort. The best estimates come from Bird Is. (Fig A1-7, A1-8) where all eggs are counted on the island as of Jan.31, and then the total number is adjusted to account for observed egg loss since the beginning of laying, data which are collected daily in a subsample of the colony. Fledglings are counted as chicks banded in October, minus chicks found dead before leaving the colony (Croxall et al., 1990). Hatching success, the proportion of eggs laid that hatch, is more easily determined in the field, and sometimes represents the best estimate of breeding success when time on the colony is limited, for example on Marion Is. (Watkins, 1987). Breeding success estimates for Antipodes and Adams Islands were estimated for a much shorter time (Walker et al., 1999; Walker and Elliot 1999; Appendix 1).

Rates of Population Change have been calculated from long-term data at only three colonies for the *Diomedea* species listed above. Croxall et al. (1990) report a steady decline of 1.0% in the Bird Is. population from 1961 to 1989. Weimerskirch et al. (1997) report a rate of population increase on Possession Is. of 0.93 for 1970-1976, 0.986 for 1977-1985, and 1.031 for 1986-1994. de la Mare and Kerry (1994) report a strong increase in the small Macquarie Is. population during 1956-1966, then an annual decline of 8.1% until 1981. Though data are limited, the Kerguelen population appears to show a similar pattern to Possession and Macquarie Is., with an apparent steep decline during the 1970s and a slow recovery since the mid-1980s.

### *Population model*

Here we present a general description of the model. The model parameters, variables and equations can be found in Appendix I. We favored a full age-structured model over a simpler model because the dynamics of age-structure can produce population behaviors that obscure true trends in population change. This is particularly important when dealing with long-lived species with delayed reproduction, such as albatross. We did not include density-dependent effects. Non-linear formulations would require us to make strong assumptions about how and when density-dependence occurs, information that is almost nonexistent for albatross populations. A linear model, we believe, is able to capture the major features of population behavior and sampling variability as they pertain to trend reconstruction.

The model is a mixed age and stage-structured model. The age structure of juveniles is explicitly represented. But once individuals mature they “recruit” into an aggregated adult class. The probability of juvenile birds “recruiting” to the adult population is controlled by a maturation function, which represents the probability that an immature bird that reaches a particular age matures and breeds. The juvenile ages are divided in two stages. This allows survival rates for young and old juveniles to be different. Since juveniles may return to the colony by 4-5 years of age yet not breed for another few years (Croxall et al., 1990; Weimerskirch et al., 1997), it is possible to calculate different survival rates for young and old juveniles.

Breeding frequency in biennially breeding albatrosses is contingent upon recent breeding success, as well as individual variation. Successful breeders next breed in 2 to 4 years and unsuccessful breeders next breed in 1 to 3 years (Croxall et al., 1990). Therefore, in our model we have made annual breeding frequency into a series of probabilities based on most recent breeding success, with each breeding frequency in each breeding success category having a different probability. The number of breeders and non-breeders in each year are modeled as functions of breeders present in previous years. Therefore, unlike more typical linear models (Caswell 1989) where transitions between all stages in a time period are explicitly modeled, here we add time delays of different duration, and calculate some stages as complements to numbers calculated for other stages. This approach allowed us to keep all the structure arising from the complex reproductive schedule of albatrosses. One of the few alternatives to this approach would have been a much more complex individual based model (De Angelis et al. 1992).

To acknowledge the difference between natural survival and induced additional mortality, the model deals with two sets of rates. Baseline parameters that represent baseline survival from natural causes are used for each of the two juvenile stages and for adults. Additional induced mortality can be added to any of these three stages and for any time window within the time horizon simulated by the model.

We included two sources of natural variability in the population. The first consists of annual variability in breeding success. The second one consists of variability in annual survival rate of juveniles and adults. Sampling variability, on the other hand, is included as “errors” added to counts of breeding pairs, the amount of error being controlled by a model parameter.

### ***Trend analysis***

Three methods were used to estimate trends in the simulated data:

- a) Linear regression approach
- b) Maximum likelihood estimation
- c) A randomization test based on resampling the observed time series

#### Linear regression approach

This procedure consists of regressing the log of the observed number of breeders against time (Gerrodette, 1987; Thomas, 1996). Assuming an exponential decay:

$$B_t = B_o e^{r t + e}$$

$$\ln(B_t) = \ln(B_o) + r t + e \quad ; \quad e \approx N(0, s^2)$$

The slope of the linear regression provides an estimate of the instantaneous rate of change (negative for a declining population). A standard one-sided *t*-test is performed to test for the null hypothesis  $H_0: r \geq 0$ .

#### Maximum likelihood approach (MLE)

Instead of regressing numbers against time, this approach tests the hypothesis  $H_0: \ln(N(t+\Delta)/N(t)) = 0$ , concentrating on changes of consecutive population size estimates. Delta,  $\Delta$ , indicates the time interval between observations, which can be different than one and can change



over time. The approach is analogous to regressing  $\ln(N(t+\Delta)/N(t))/\sqrt{\Delta}$  on  $\sqrt{\Delta}$  without an intercept, estimating the trend from the slope and performing a  $t$ -test on it.

Like the regression approach, the MLE approach assumes an exponential decline and is derived from a diffusion approximation to population trends (Dennis et al. 1991). Underlying both methods, however, lie different assumptions about the error structure in the observations. The regression approach assumes that the trend is deterministic and that the residual variability is uncorrelated between years (*i.e.*, observation error model, Hilborn and Mangel, 1997). Conversely, the MLE approach is only exact if there are no observation errors. Further, it assumes that the residuals of the regression of  $\ln(N(t+\Delta)/N(t))/\sqrt{\Delta}$  on  $\sqrt{\Delta}$  are uncorrelated. The assumptions of either method are not met by biennially breeding albatross species. The particular breeding behavior of these species creates a strong lag-1 negative autocorrelation in counts of breeding pairs. This structure obscures the decline, reducing the power of both methods to detect a trend. As will be shown later, this is particularly critical for the MLE method, which is rendered useless by the violation of its assumptions.

#### Randomization test

This statistical method was chosen because of the need to accommodate better the structure in the data, thereby increasing the power to detect a trend. It consists of what is usually referred to as a *randomization test* (Manly, 1991), a procedure similar to a *moving block bootstrap* (Efron and Tibshirani, 1993). The approach is similar to that previously applied to the South Georgia albatross data (section A *preliminary analysis*). As before, data sets are “reconstructed” by randomly sampling the data, but instead of sampling residuals, here we sampled with replacement sets of 3 consecutive years of data from the “observed” (in our case simulated) time series of counts of breeding pairs. These windows are patched together to build a time series of length equal to the specified sampling period. This process is repeated several times and for each reconstructed data set a slope is estimated with the linear regression approach. Because the data are resampled by “chunks”, the reconstructed data sets will keep most of the autocorrelation structure observed in the original data. But, since the windows are reordered at random, the reconstructed data sets should be “free of trend”. Therefore, the distribution of slopes that arises from the generated data sets provides a distribution of the parameter of interest under the null hypothesis (*i.e.* no trend). This

distribution is therefore used as a reference distribution against which the trend in the original data set is contrasted. When the magnitude of the observed trend is “rare”, as suggested by the reference distribution, evidence exists for a real trend. On the other hand, finding that the magnitude of the observed trend is “common” in the reference distribution, provides an indication that there is no trend in the data. The area under the reference distribution beyond the observed value provides a  $p$ -value, and can be used to decide whether to give credence or not to the hypothesis that a trend actually exists.

### *Implementation*

The model is implemented as an Excel spreadsheet which uses Visual Basic macros to iteratively simulate multiple realizations of the same process (Appendix 3). The macros also generate random numbers faster and more efficiently than Excel built in capabilities. All population parameters in the model can be controlled, as well as sampling parameters controlling when sampling is performed, for how long and how frequently (annually, biannually, etc.), and also the amount of sampling error involved. The output consists of a graphic time series of population trajectory and estimated population counts. Quantitative measures of the actual trend and population variability are calculated together with estimated trend and significance tests for the sampling window.

The spreadsheet can be run in two modes. In the first one a one-time projection is done over a 200 year period and tests are performed for a selected sampling window. The second mode allows to perform a Monte Carlo simulation based on generating multiple population trajectories and sampling windows of different lengths. The output of the simulations are written to an ASCII file. For each replicate in the simulation, the average of the actual population trend, its variability and the results of statistical tests for different window sizes are recorded.

### *Scenarios simulated in this report*

In order to generate results, we selected a number of scenarios, which combine different trend sizes, different amounts of sampling error, and varying sampling periods and frequencies. We also tried cases where the trend was stabilized when sampling started and cases where the sampling started simultaneously with a reduction in survival of young juveniles, producing a declining trend

in an otherwise non-growing, non-declining population. We chose this last perturbation because changes in young juvenile survival are expected to be detected with the greatest delay in adult breeding counts. The scenarios tried were:

**A) Stabilized trends, sampling every year**

- a1) No sampling error (CV=0), trends of 0%, 1% and 2% per year
- a2) Moderate sampling error (CV=5%), trends of 1% and 2% per year
- a3) Significant sampling errors (CV=10%), trends of 1% and 2% per year

**B) Stabilized trends, sampling every other year**

- b1) Moderate sampling errors, trends of 1% and 2% per year

**C) Transient behavior, sampling every year**

- c1) Moderate sampling errors, trends of 1%, and 2% per year

Appendix 4 contains the parameter values used to simulate each of these different behaviors.

## **Results**

### ***Model validation***

In order to evaluate the performance of the model, we generated a 200 year time series of breeding populations (scenario “a1” with a 1% trend ) and compared it to the data available for South Georgia. Figure 4 shows three representative subsets of 17 years of the simulated data, which can be visually contrasted to the actual data (Figure 1). The residual variance in the produced time series is very similar to that in the SG data (Mean Square Error: 0.0036 in South Georgia, 0.003 for simulations with no sampling error).

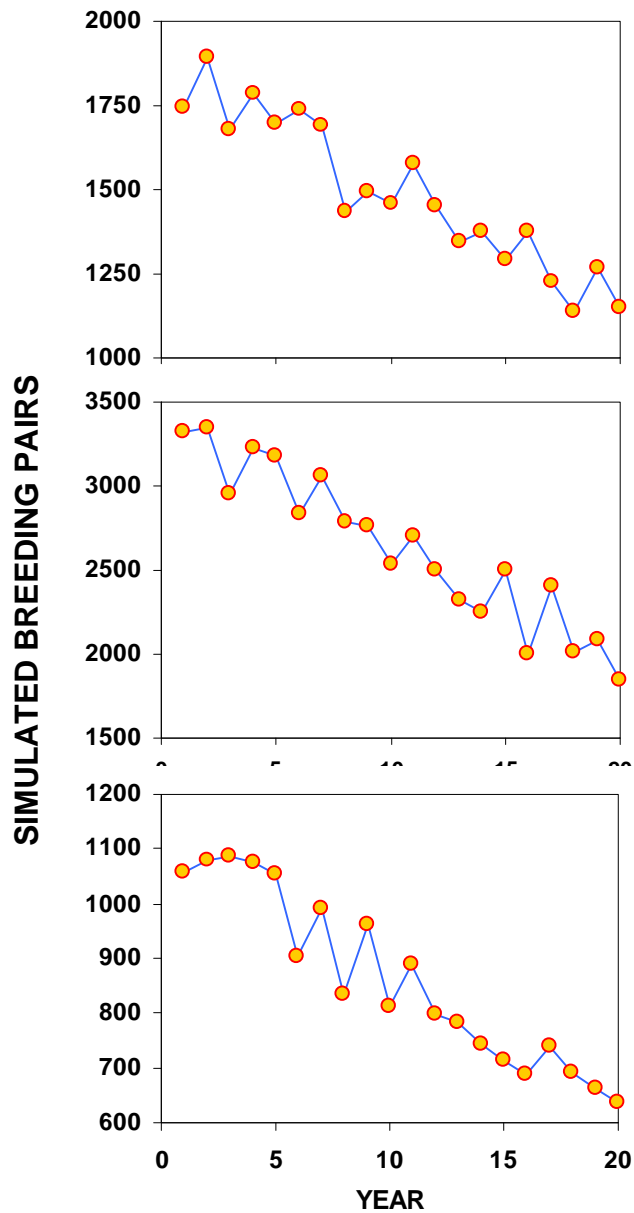


Figure 4. Three subsets generated with the model with a trend of 1% per year.

Another feature that is well reproduced by the model is the autocorrelation structure characteristic of South Georgia data. Figure 5 (upper panel) shows the residuals of the fit to the South Georgia data (see Figure 1), which show alternating low and high values. This pattern arises from the fact that a good reproductive year (high breeding success) is followed by low colony

attendance next year due to a large number of birds skipping reproduction. This produces a significant lag-1 autocorrelation, as shown in Figure 5 (lower panel). This temporal correlation structure in the data is well reproduced by the model. Figure 6 shows the distribution of lag-1 autocorrelation coefficients in sets of 17 consecutive years (the length of SG data) simulated by the model. The level of the autocorrelation observed in the South Georgia data set emerge as a very likely occurrence in the population behavior simulated by our model.

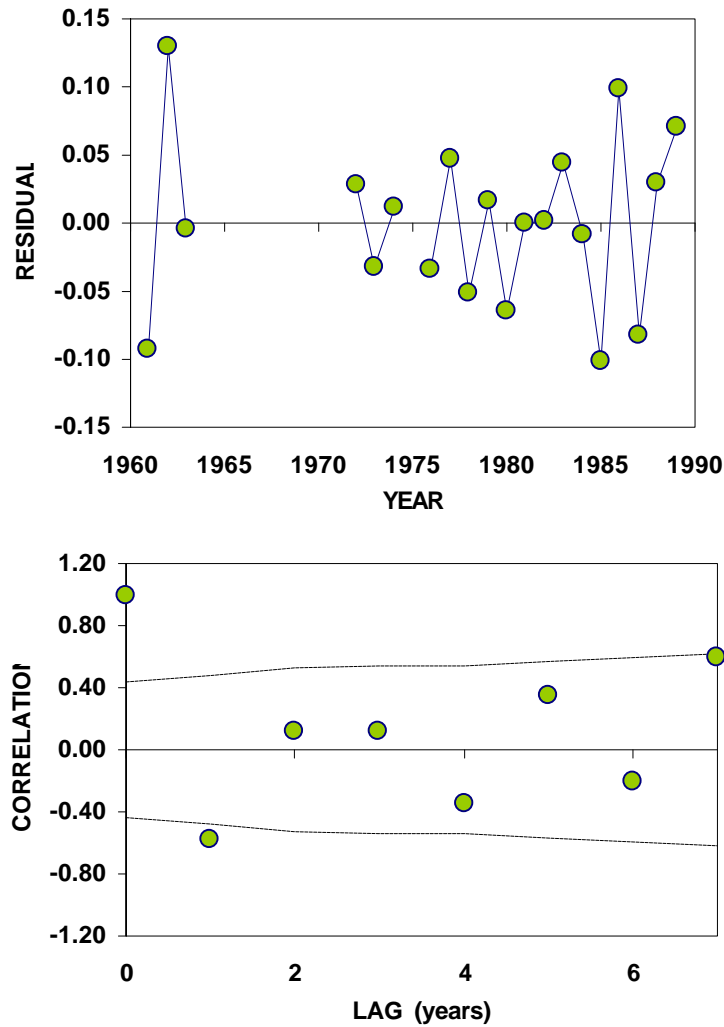


Figure 5. Upper panel: residuals from the fit of an exponential decay model to the South Georgia data. Lower panel: autocorrelogram of South Georgia data.

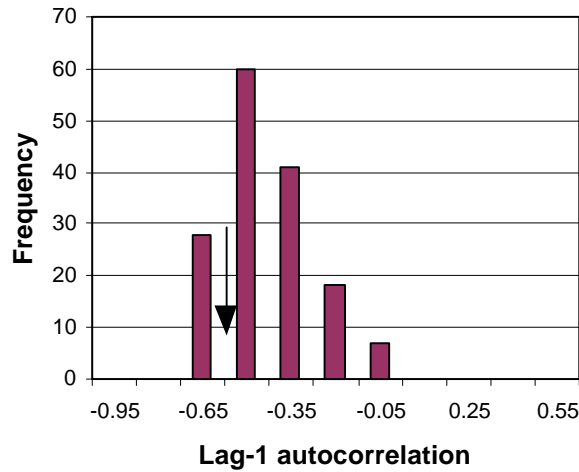


Figure 6. Frequency distribution of lag-1 autocorrelation for 154 time frames of 17 years of breeding populations simulated by the model. The arrow indicates the lag-1 autocorrelation in the data from South Georgia.

The size and shape of the variability, in addition to trend size, critically affect the power to detect a trend. Our model and selected parameters mimic extremely well both characteristics of data from real populations.

***Power analysis for trends in counts of breeding pairs***

As shown in Figure 7, the MLE approach failed completely to detect simulated trends. This complete failure is due to the particular structure characteristic of breeding counts of biennial species discussed in the previous section. The basic information used by the MLE approach to estimate the trend and its significance is the ratio of population sizes in consecutive years, which is strongly masked by the lag-1 negative correlation in the data. All simulations performed showed the same failure of this method and we, therefore, do not report specific results and conclude that in its current format this method is not suitable for analyzing this type of data.

The regression approach provided more satisfactory results. Figure 8 shows the power curves derived for the case of stabilized trends, when sampling is performed every year (such as those depicted in Figure 4), for trend sizes of 0, 1 and 2%. The three panels correspond to simulations that include different levels of sampling error. The results for runs without sampling error (upper panel) are almost identical to those derived by the resampling procedure of South

Georgia data (Figure 2, upper panel). This is not surprising, because the basic biological parameters we used in the simulations were derived mostly from South Georgia studies. Moreover, it provided additional support, albeit more indirect, for the model we developed. As before, these results indicate that to attain an 80% probability of detecting a trend of 1% a year, 15 years of data are required. Ten years of data provide a low power (about 25%) to detect a trend of 1% and much higher power is attained for the same amount of data if the trend is 2% per year (about 80%).

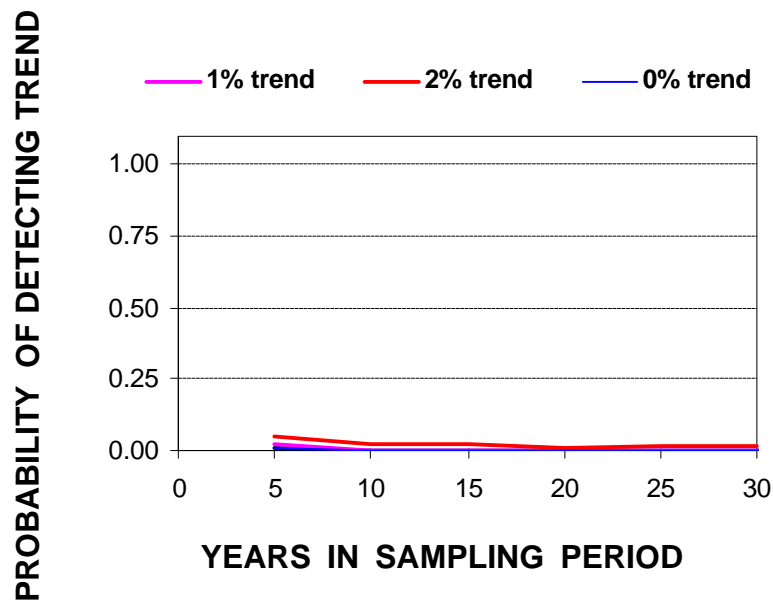


Figure 7. Power curves for the detection of stabilized trends of different size for the Maximum likelihood approach as a function of years in the sampling period. Counts were simulated to take place every year.

Adding sampling variability (middle and lower panel) results in a significant deterioration of power to detect the trend. What is here called a “moderate” error consists of simulating observed values that 50% of the time are less than 10% off the true value, while “large” errors produce 50% of the observations to be less than 20% off the true value. Simulations with moderate errors (middle panel in Figure 8) indicate that 20 years are required for a 50% chance of detecting a trend of 1% and the power increases to 100% if the trend is of 2% for the same number of years. Large errors extend the time required for a 50% chance to detect trends of 1% to 27 years.

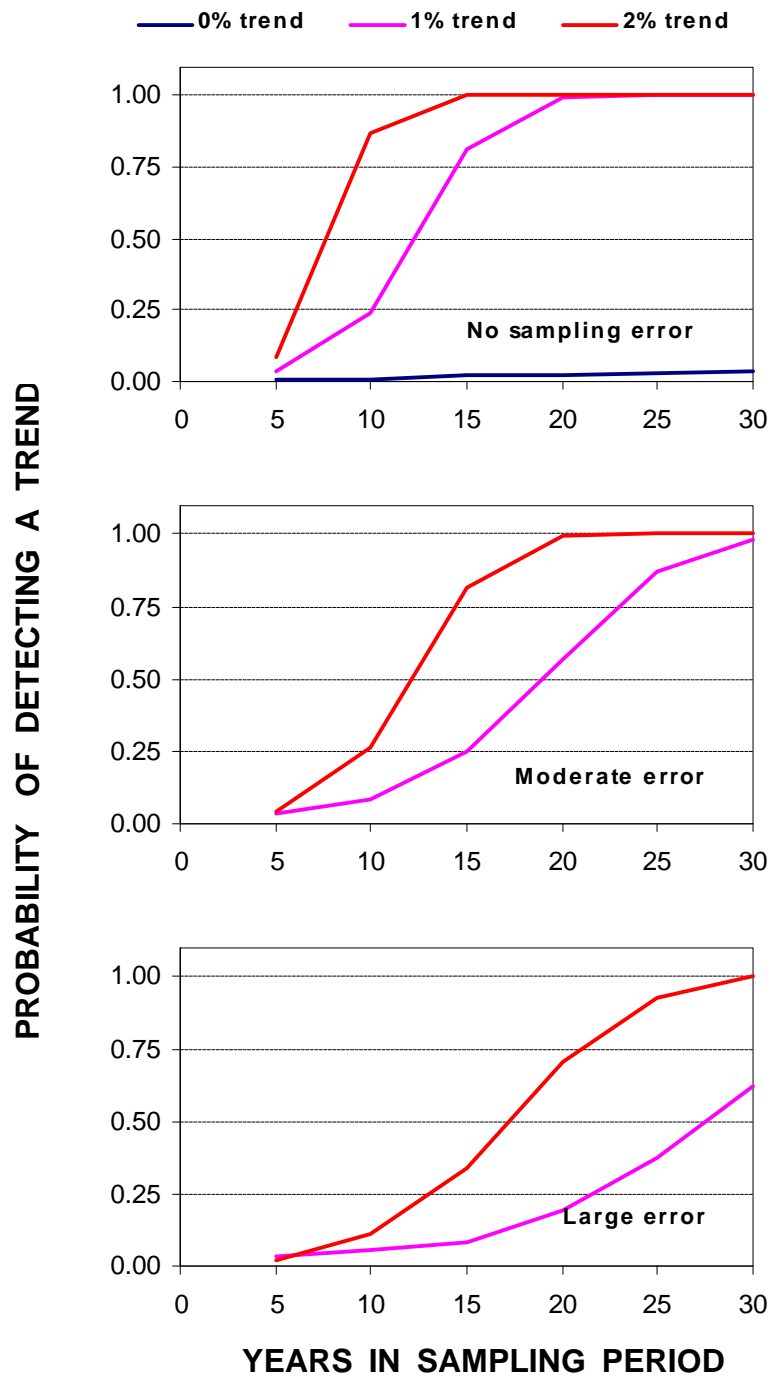


Figure 8. Power curves for the detection of stabilized trends of different sizes, and for increasing amounts of sampling error, as a function of years in the sampling period. Counts were simulated to take place every year.

For moderate sampling error and trends of 1%, recording population size every other year (Figure 9) will increase the power of ten population counts (taken in this case over a 20-year



period) from less than 10% to over 25%. Analogously, the power provided by 15 population counts increases from 25% to 75% when those counts are taken over a 30-year period.

One of the most serious problems with detecting trends based on counting adults is that the effects of incidental mortality on young stages will show up with a delay of years. This delay will be more serious as longevity increases. The power curves for a trend produced by a decrease in the survival of young juveniles that starts when the censuses begin are given in Figure 10 (compare with the results for stabilized trends: Figure 8 middle panel). In this case, the power of a short time series (<15 years) to detect trends of both 1 and 2% is seriously undermined. Longer time series provide enough years for the trend to be manifested and 30 years of data provide as much power as for the case with stabilized trend.

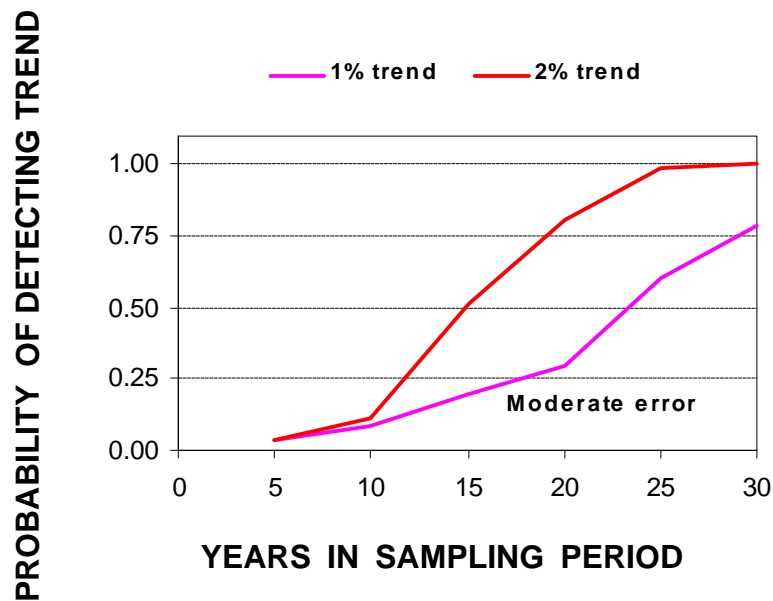


Figure 9. Power curves for the detection of stabilized trends of different size and for moderate level of sampling error with the regression approach. Counts were simulated to take place every other year.

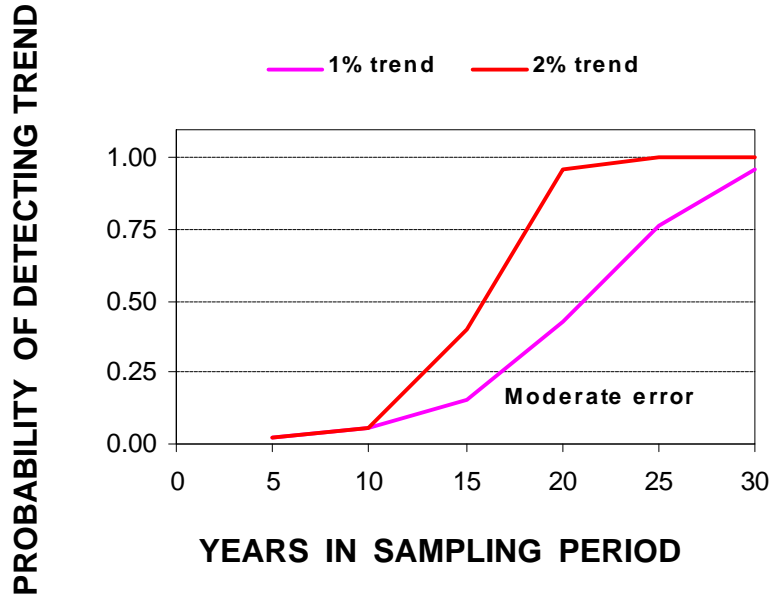


Figure 10. Power curves for the detection of transient trends of different sizes and for moderate levels of sampling error. Counts were simulated to take place every year. Transient trends were simulated by a reduction in survival of young juveniles that was timed to begin when sampling began in an otherwise non-growing, non-declining population. This produces a delay in the materialization of the trend.

In summary, when the linear regression analysis of counts of breeding pairs is used, 10 years of data provide very low power to detect trends of 1%. This is true even for the case in which there is no sampling error. The power of 10 years of data for detecting trends of 2% is significant (>75%) only if no sampling errors exist. As sampling error increases, at least 20 years of data (and up to 30 years for the scenarios we chose) are required to provide powers of over 50% to detect trends of -1%.

***Power analysis with resampling methods***

The randomization method based on resampling 3-year data blocks provided higher power than the simple regression approach (Table 1). This improvement is particularly pronounced for simulations without sampling error, where the power to detect trends of 1% with 12 years of data increased by 45%. This method also produced higher power for simulations with moderate sampling error, although the gain was not as large.

Table 1. Comparison of power provided by simple regression and by randomization method (resampling of 3-year time frames) for particular scenarios and when 12 consecutive years of counts are available.

Counts	Trend	Sampling error	Power of regression approach	Power of resampling approach
12	0%	No error	0.01	0.07
12	1%	No error	0.47	0.68
12	1%	Moderate	0.15	0.18

***Other indices of population size***

As shown before, the most critical problem jeopardizing the ability of statistical methods to detect a trend from breeding surveys is the negative autocorrelation characteristic of these data. An additional problem with the use of breeding counts as an index of population health is that it will show effects on juveniles only after significant delay. An ideal index of population health would be based on population stages whose presence at identified sites are more predictable. The younger the birds counted, the faster the index will show signs of decline.

As mentioned before, older albatross juveniles starting from about 4 or 5 years of age visit the colony for a number of years prior to actively breeding. They do this to practice breeding rituals in preparation for obtaining a long-term mate and nest site. In order to explore the value of counts of these younger stages as indices of population health, we repeated the simulations of scenario A, but this time we applied the regression approach to counts of juveniles, to which we added different levels of sampling error. The basic biological variability that affects juvenile numbers in our simulations is cohort strength, given by inter-annual variability in breeding success and the basic binomial error associated to the mortality process. This variability does not include inter-annual changes in the tendency of juveniles to visit the colony, for which we do not possess information. Therefore, “sampling error” should be regarded here as the overall variability associated with counting extant juveniles in the colony due both to colony attendance and counting errors. Figure 12 shows the power curves for the regression analysis of counts of old juveniles (>5 years) when different levels of this “sampling” error are added to the numbers in the populations. If no errors exist, a scenario consistent with invariable colony attendance by juveniles, the power to detect a trend of 1% is very high (80%).

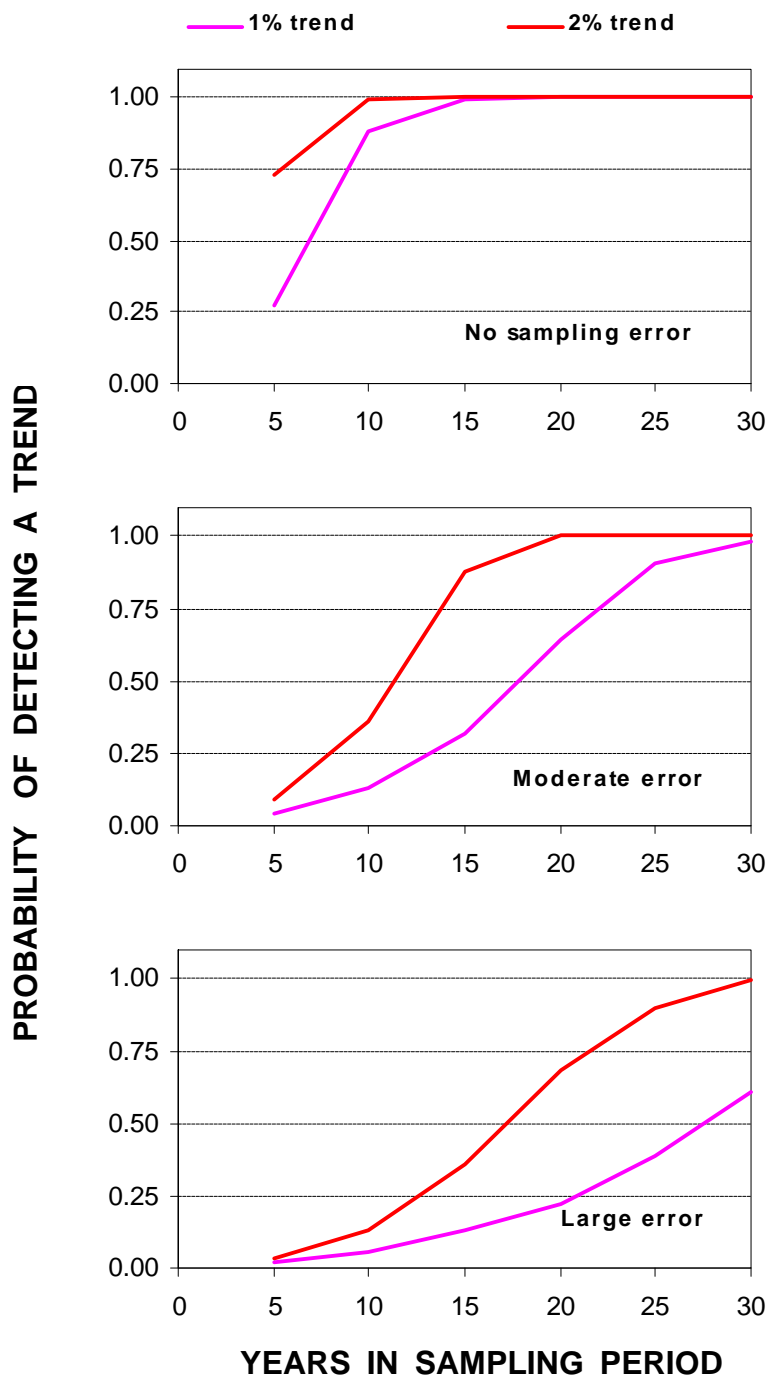


Figure 12. Power curves for the detection of stabilized trends of different sizes in counts of older juveniles (age 5+) and for increasing amounts of sampling error. Counts were simulated to take place every year.

As variability in the counts increases, the power decreases. In fact, for a moderate error level, the power provided by counting juveniles is marginally higher than that provided by counting breeding

pairs. This result suggests that errors added to the basic biological variability have an overriding effect on the power to detect trends.

## **Discussion**

Wandering Albatross populations on sub-antarctic islands have been experiencing significant population declines over the last three decades (Croxall et al., 1990; Weimerskirch et al., 1997). There is evidence that incidental by-catch mortality of adults and juveniles by long-line fisheries may be contributing to these declines (Brothers, 1991). Breeding albatross populations on Adams and Antipodes Islands, New Zealand, constitute a large proportion of the total number of great albatrosses found in the southern oceans, yet systematic monitoring of these populations began only in the early 1990s (Walker et al., 1995; Clark et al., 1995). Appropriate management of these albatross populations, in part by implementing strategies to reduce long-line bycatch, requires information about population trends, and the demographic factors that most strongly influence those trends. Because of “demographic momentum”, estimating population level trends in long-lived species such as albatrosses requires many years of monitoring data. Wandering Albatrosses add to the problem of detecting population level trends because they are biennial breeders. Thus, annual counts of breeding pairs can fluctuate significantly between successive years, even when population size remains stable. We have developed a series of approaches intended to maximize the likelihood of detecting population level changes in biennially breeding albatrosses.

We were able to develop a population model for biennially breeding albatrosses, while keeping it general enough to accommodate a broad range of life history characteristics, including annual breeding. Using this model and a select set of parameters we were able to derive some general conclusions about how many data points are needed to detect population level trends of different magnitudes. We also were able to explore the effects of sampling error on our ability to detect trends and scrutinize the performance of alternative statistical methods.

How much statistical power do counts of breeding pairs of biennially breeding albatross populations over a limited number of years provide for detecting population level trends? This was the main question driving the development of this project. In general, 10 years of data provided very low power to detect trends of 1% a year. The life-history driven, inter-annual variability in the size of the breeding population, together with sampling errors and inherent individual and environmental variation, requires longer time series for trend detection. Our results suggest that at

least 15 years of data will be required to detect declining trends of 1% per year in populations of biennially breeding albatross species when sampling errors are insignificant. Power deteriorates quickly as sampling errors increase. Therefore, it is critical that sampling protocols be of the highest quality, and produce accurate and precise estimates.

The main factor producing biological variability in our simulations was biennial breeding. This provides an irreducible source of variability in counts of breeding pairs, which diminished the power of the regression approach and caused the Maximum Likelihood Estimation approach to fail completely. The only way to overcome the limitations of the MLE approach would be to derive a likelihood that accounts for the complexity of intermittent breeding. This may not be straightforward because the diffusion approximation derived by Dennis et al. (1991) is based on a random walk model between consecutive years. In any case, it lies beyond the objectives of this project.

Based on our results we identified three avenues to circumvent the low power provided by traditional statistical methods as applied to biennial species. The first and most direct option was to use statistical methods that can better accommodate the autocorrelation typical of these data. The randomization approach we implemented, based on resampling blocks of data, worked significantly better than the simple regression approach. The next step would be to perform a more detailed analysis of optimal window size and explore how the best choice depends on the breeding schedule of the species. The downside of randomization methods is that they are computer intensive, but they can be implemented with relatively simple tools (spreadsheets and macros).

The second option we used to increase the power to detect population trends was to look for alternative indices of abundance. As an alternative to counts of breeding pairs, counts of old juveniles at the colonies hold the greatest promise because they are not expected to display the inherent annual fluctuations that are characteristic of counts of breeding adults. The value of juvenile counts as an index of abundance will depend on how consistently they return to the colony, and how accurately they can be counted at the colony.

A third option to increase the power would be to change the alpha level used in tests of significance when the simple regression approach is used. Negative autocorrelation decreases the power to detect trends and using a higher alpha level would decrease the burden of proof put on time series counts of breeding pairs. We do not favor this approach because we believe that it will be difficult to determine objectively an alpha level that will satisfy both the management need to

detect trends early with the need for an objective diagnosis. Another reason is that we believe that conservation efforts and contingent management plans should go beyond the simple diagnostic provided by a significance test and be grounded on decision analysis protocols that accommodate an explicit account of the data available, the uncertainties associated with these data, and the management options at hand. More sophisticated statistical analysis tools or the use of auxiliary abundance indices provide a more direct way to estimate trends given the particular structure of monitoring data available for biennially breeding albatross species.

We used the framework developed here to derive general conclusions about trend detection and to generate needed background information about the value of different statistical techniques. Nevertheless, these tools also are suitable for exploring the dynamics of particular populations of interest.

There is one caveat associated with our choice of sampling error. The lognormal error used is simple to program and easy to generate on a computer. Furthermore, it has the advantage of being bounded to positive values and its magnitude is controlled by a single parameter that provides an intuitive account of the level of variation (% error). It produces a constant coefficient of variation, meaning that relative effects do not depend on the expected value or on the particular level of the population. These attributes led us to choose this distribution over alternative ones. It must be understood, however, that a lognormal distribution could conceivably produce extreme counts more often than alternative error structures, such as binomial errors. Therefore, it is extremely important for users applying this approach to analyze the power to detect trends, make an *a priori* judgement of the nature and magnitude of the sampling error expected in their particular system or, alternatively, repeat the analysis for different candidate error structures.

## References

- Bergin, A. 1997. Albatross and longlining – managing seabird bycatch. *Marine Policy* 21:63-72.
- Brothers, N. 1991. Albatross mortality and associated bait loss in the Japanese longline fishery in the Southern Ocean. *Biol. Conserv.* 55:255-268.
- Caswell, H. 1989. *Matrix Population Models: construction, analysis, and interpretation*. Sinauer Associates, Inc. Sunderland.
- Clark, G., J. Amey and G. McAllister. 1995. Unexpectedly large number of Wandering Albatrosses (*Diomedea exulans*) breeding on Antipodes Island, New Zealand. *Notornis* 42:42-46.
- Croxall, J.P., P. Rothery, S.P.C. Pickering and P.A. Prince. 1990. Reproductive performance, recruitment and survival of wandering albatrosses *Diomedea exulans* at Bird Island, South Georgia. *J. of Animal Ecology* 59:775-796.
- De Angelis, D.L. y L. Gross (Ed.). 1992. *Individual Based Models and Approaches in Ecology: Populations, Communities and Ecosystems*. Chapman and Hall, New York.
- de la Mare, W.K. and K.R. Kerry. 1994. Population dynamics of the wandering albatross (*Diomedea exulans*) on Macquarie Island and the effects of mortality from longline fishing. *Polar Biol.* 14:231-241.
- Dennis, B., Munholland, P. and J.M. Scott. 1991. Estimation of Growth and Extinction parameters for endangered species. *Ecological Monographs* 61:115-143.
- Efron, B. and R.J. Tibshirani. 1993. *An introduction to the Bootstrap*. Monographs on Statistics and Applied Probability 57. Chapman and Hall, London.
- Gales, R. 1998. Albatross populations: status and threats. Chapter 3 in Robertson & Gales [eds] *Albatross Biology and Conservation*. Surrey Beatty & Sons, New South Wales, Australia.
- Gerrodette, T. 1987. Power Analysis for detecting trends. *Ecology* 68:1364-1372.
- Hilborn, R. and M. Mangel. 1997. *The Ecological detective: confronting models with data*. Monographs in Population Biology 28. Princeton University Press, New Jersey.
- Jouventin, P. and H. Weimerskirch. 1991. Changes in the population size and demography of southern seabirds: management implications. Chapter 14. In: *Bird Population Studies: Relevance to Conservation and Management*. Perrins, Lebreton and Hirons [eds]. Oxford University Press.



- Manly, B.F.J. 1991. Randomization and Monte Carlo Methods in Biology. Chapman and Hall, London.
- Robertson, C.J.R. 1975. Report on the distribution, status and breeding biology of the Royal Albatross, Wandering Albatross and White-capped Mollymawk on the Auckland Islands. In: Preliminary Results of the Auckland Islands Expedition, 1972-1973. Dept. of Lands and Survey, Wellington, N.Z.
- Robertson, C.J.R. and Nunn, G.B. 1998. Towards a new taxonomy for albatrosses. Chapter 2 in Robertson & Gales [eds] Albatross Biology and Conservation. Surrey Beatty & Sons, New South Wales, Australia.
- Thomas. 1996. Monitoring long-term population change: why are there so many analysis methods? Ecology 77:49-58.
- Walker, K., P. Kilks, G. Elliott and J.-C. Stahl, 1991. Wandering Albatross on Adams Island, February, 1991. New Zealand Department of Conservation, Science & Research Internal Report No. 109.
- Walker, K., G. Elliott, A. Davis and P. McClelland, 1995. Wandering Albatross on Adams Island: Census, Nesting Data and Body Measurements, February, 1993. Science & Research Series No. 78. Dept. of Conservation, Wellington, N.Z.
- Walker, K., G. Elliot, S. Hamilton and A. Wiltshire. 1999. The impact of fisheries bycatch on Antipodean Wandering Albatross. Report on the 1998/99 field season. New Zealand department of Conservation, Wellington, N.Z. Draft report.
- Walker, K. and G. Elliot. 1999. The impact of fisheries bycatch on Gibson's Wandering Albatross. Report on the 1998/99 field season. New Zealand department of Conservation, Wellington, N.Z. Draft report.
- Watkins, B.P. 1987. Population sizes of king, rockhopper and macaroni penguins and wandering albatrosses at the Prince Edward Islands and Gough Island, 1951-1986. S. Afr.J.Antarct. Res. 17(2):155-162.
- Weimerskirch, H., N. Brothers and P. Jouventin. 1997. Population dynamics of Wandering Albatross, *Diomedea exulans*, and Amsterdam Albatross, *D. amsterdamensis*, in the Indian Ocean and their relationships with long-line fisheries: conservation implications. Biol. Cons. 79:257-270.

## APPENDIX 1: ALBATROSS DATA

### 1. Number of Breeding Pairs

Due to the logistical difficulty of monitoring most albatross breeding colonies, counts of breeding pairs are limited. Complete counts of breeding pairs are listed by year in Table A1-1.

The data in Table A1-1 are represented graphically in Fig. A1-1 in order to illustrate a) differences in the magnitude of population means and variances, b) differences in the frequency of sampling, and c) differences in population trends between wild albatross populations.

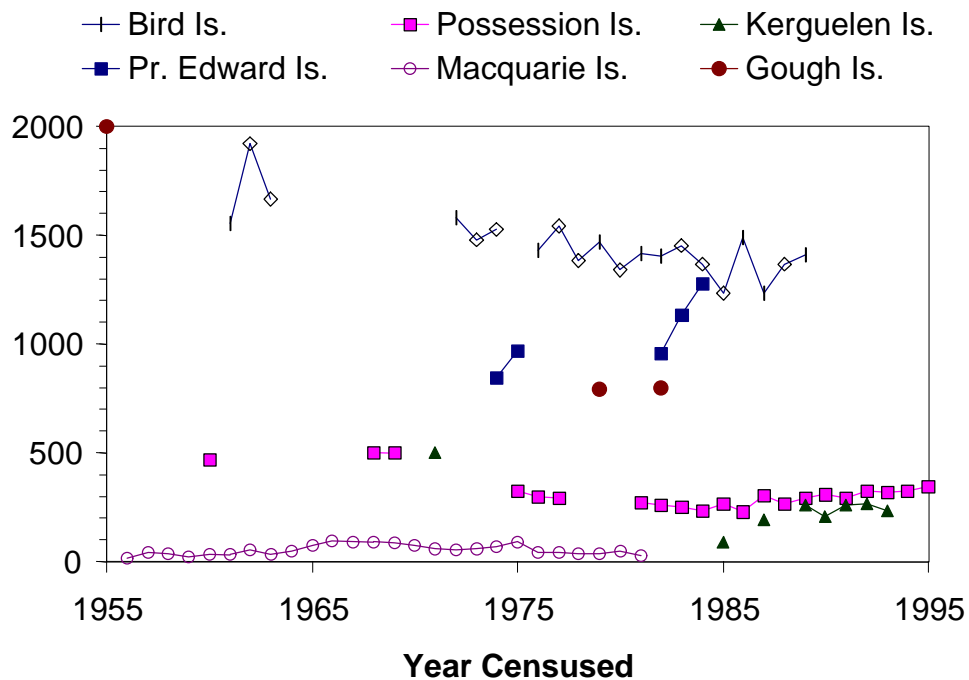


Fig. A1-1: Annual counts of breeding pairs for *Diomedea dabbenena* (Gough Is. only) and *D. exulans*. Counts for *D. gibsoni* and *D. antipodensis* were too large to include in this figure. Lines designate continuous annual counts.

Note that data for Adams and Antipodes Islands (with populations of about 4000 and 4500 breeding pairs annually) are not included in Fig. A1-1. The most recent whole population censuses indicate that an average of 5136 breeding pairs visited Antipodes Island (*D. antipodensis*) in the period 1994-97 (Walker et al. 1999) and an average of 5831 visited Adams Island (*D. gibsoni*) in the period 1991-97 (Walker and Elliot 1999). Whole population censuses have been discontinued at

both sites; counts are now performed of only a representative portion of the island and are intended to monitor population change rather than assess population size.

**Table A1-1:** Annual counts of breeding pairs of four *Diomedea* species, listed by island.

YEAR	<i>Diomedea exulans</i>						<i>Diomedea dabbenena</i>	<i>Diomedea gibsoni</i>	<i>Diomedea antipodensis</i>
	Bird <sup>a</sup>	Possession <sup>b</sup>	Kerguelen <sup>b</sup>	Marion <sup>c</sup>	Prince Edward <sup>c</sup>	Macquarie <sup>d</sup>	Gough <sup>e</sup>	Adams <sup>f</sup>	Antipodes <sup>g</sup>
1951				700					
1955							~2000		
1956						17			
1957						43			
1958						35			
1959						19			
1960		470				34			
1961	1554					34			
1962	1922					51			
1963	1666					32			
1964						50			
1965						76			
1966						97			
1967						90			
1968		500				92			
1969		500				87			~1000
1970						76			
1971			500			58			
1972	1582					55		~7000	
1973	1477					56			
1974	1529			1557	847	70			
1975		325		1852	966	93			
1976	1433	300				40			
1977	1541	295				41			
1978	1382					36			
1979	1466					36	792		
1980	1339					49			
1981	1415	270				27			
1982	1404	260		1207	957		798		
1983	1453	250		1137	1135				
1984	1366	235		1184	1277				
1985	1232	265	90	1168					
1986	1491	230		1096					
1987	1233	305	190						
1988	1366	265							
1989	1411	295	260						
1990		310	210						
1991		295	260					4000	
1992		325	265						
1993		320	235					3874	
1994		325						?	4522
1995		345							

References

a: Croxall et al., 1990, Table 1.

b: Weimerskirch et al., 1997, Fig 2.

Note: Numbers of breeding pairs on Possession and Kerguelen estimated from figure to nearest 5 pairs.

c: Watkins, 1987, Table 5.

d: de la Mare and Kerry, 1994, Table 2.

e: Watkins, 1987, Table 6.

f: Walker et al., 1991, 1995

g: Clark et al., 1995.

?: Census taken but data not yet published, Walker et al., 1995.

## 2. Adult Survival

Annual adult survival has been estimated over many years for three sites only, Bird, Possession and Macquarie Islands. The population size at Macquarie is so small that estimated annual adult survival is highly variable and so is not included here. Time series and frequency distribution data are presented in Figs. A1-2 and A1-3.

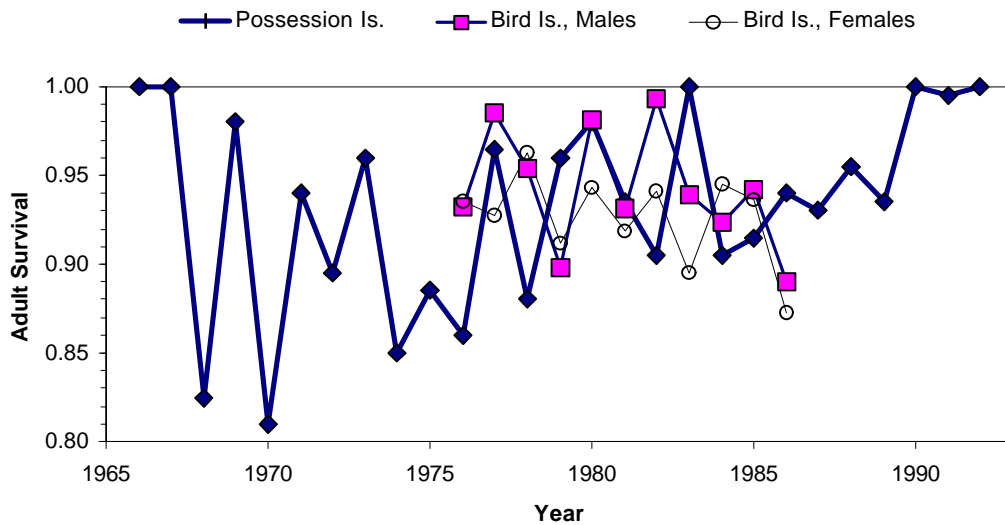


Figure A1-2: Time series of adult survival rates on Possession Is., and for both males and females on Bird Is.(data from Weimerskirch et al., 1997; Croxall et al., 1990).

Adult survival estimates for Antipodes and Adams Islands were estimated for a much shorter time. Estimated average survival for *D.antipodensis* for the period 1994-96 was 0.99 (Walker et al., 1999) and 0.97 for *D.gibsoni* for the period 1991-97 (Walker and Elliot 1999).

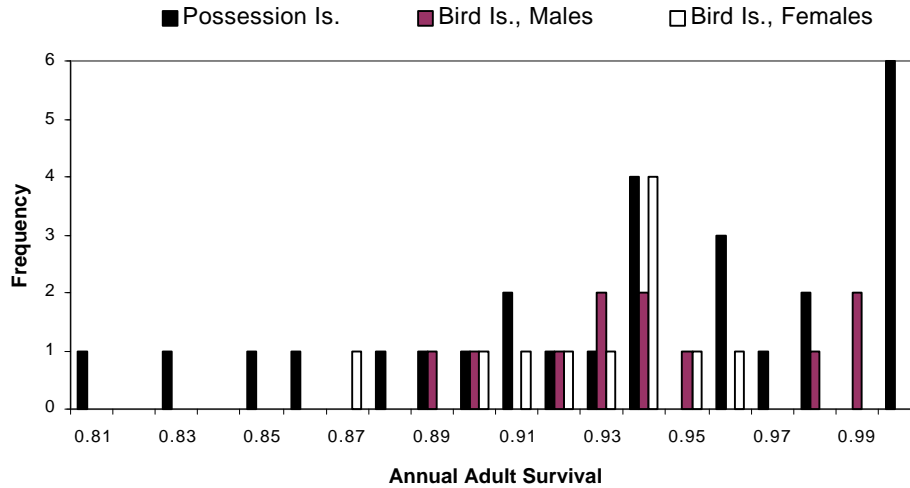


Figure A1-3: Frequency distribution of adult survival rates on Possession Is., and for both males and females on Bird Is.(data from Weimerskirch et al., 1997; Croxall et al., 1990).

### 3: Juvenile Survival and Recruitment Rates

Juvenile survival to 5 years has increased significantly on both Bird and Possession Islands (Fig. A1-4). Data on rates of recruitment from fledging to the adult breeding population are very limited, but suggest a possible decrease in recruitment from the 1960s to the 1970s (Fig. A1-5).

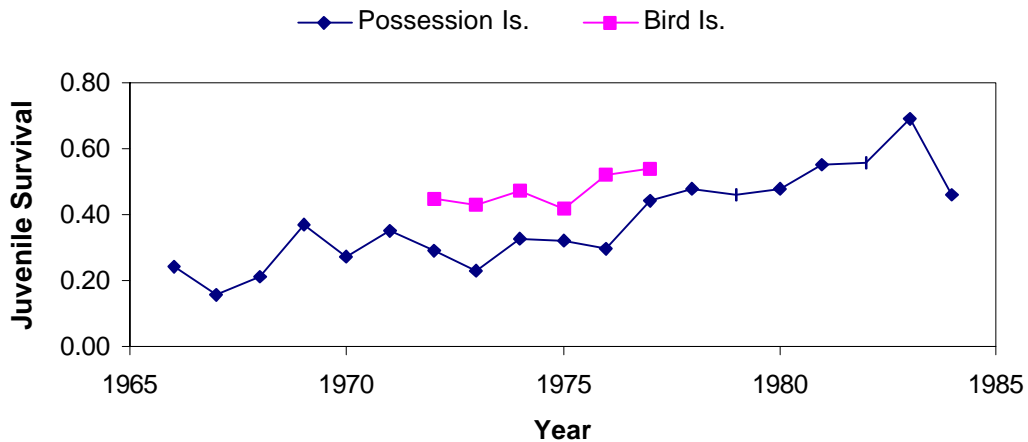


Figure A1-4: Juvenile survival to age 5 years on Bird and Possession Islands (data from Weimerskirch et al., 1997; Croxall et al., 1990).

No estimates of juvenile survival are available for *D.gibsoni* and *D.antipodensis*.

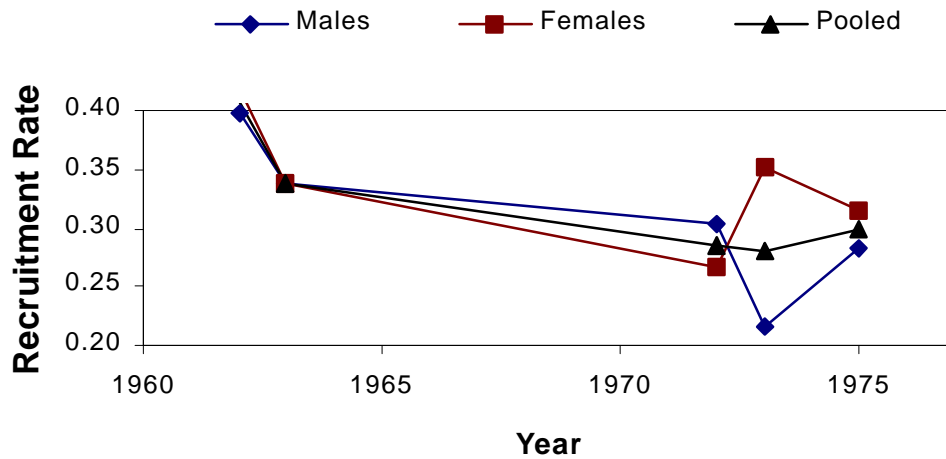


Figure A1-5: Recruitment rate to the breeding population on Bird Is. by fledging cohort (data from Croxall et al., 1990).

#### 4. Breeding Frequency

Breeding frequency data are available only for Bird Is. The probability of returning to breed two years after breeding successfully falls within roughly the same range as the probability of returning to breed one year after an unsuccessful breeding attempt (Fig A1-6a). A minority of birds return to breed 3 or 4 years after successfully breeding, else 2 or 3 years after unsuccessfully breeding (Fig A1-6b).

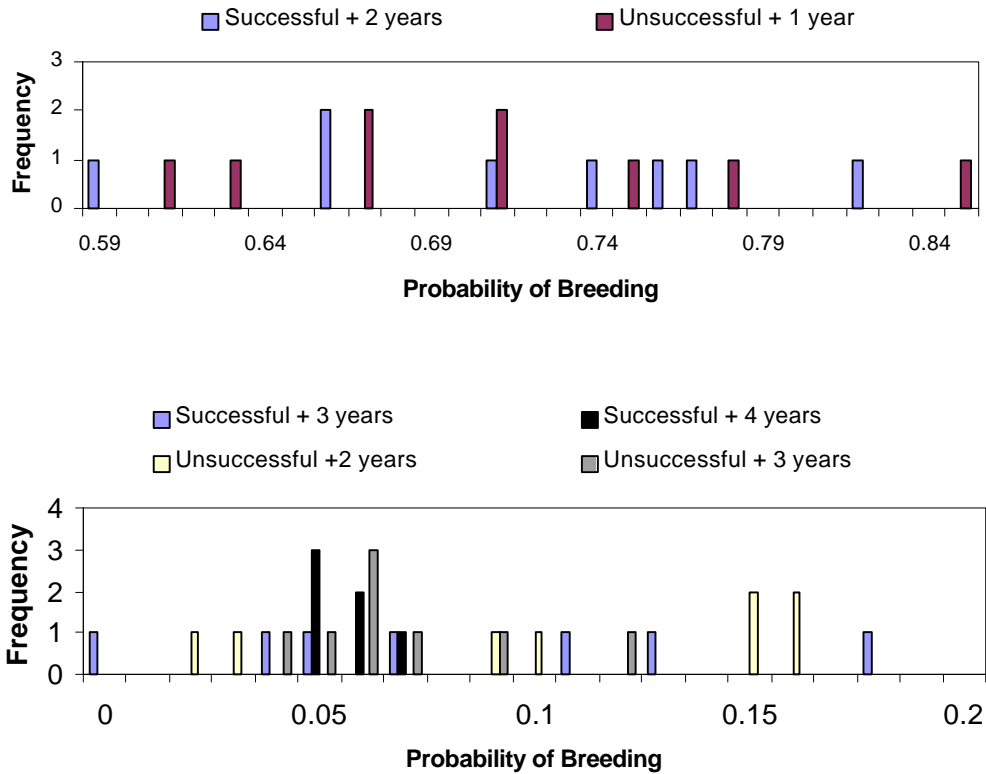


Figure A1-6a and A1-6b: Frequency distributions of the probability of breeding n years after breeding either successfully or unsuccessfully on Bird Is. between 1976 and 1985 (data from Croxall et al., 1990).

5: Breeding Success

Even though adult survival on Bird Is has decreased, breeding success has increased at about 1.2% per year, or 13% overall (Fig. A1-7). Mean breeding success is 0.64. Breeding success on Possession Is. averaged 0.68 (Fig. A1-8), and also appears to have increased significantly over 28 years.

Breeding success estimates for Antipodes and Adams Islands were estimated for a much shorter time. Estimated breeding success for *D.antipodensis* for the period 1994-98 was 0.77 (Walker et al., 1999) and 0.67 for *D.gibsoni* for the period 1991-98 (Walker and Elliot 1999).

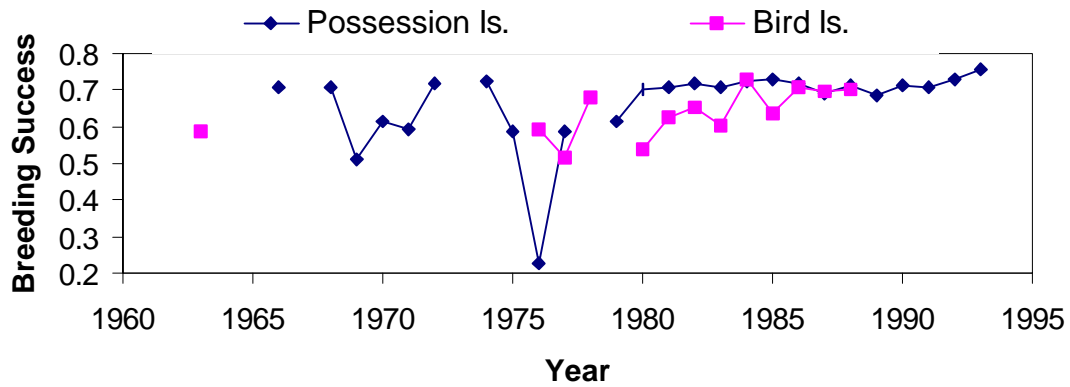


Figure A1-7: Time series of breeding success (proportion of eggs laid that produced fledged chicks) on Bird and Possession Is. (data from Weimerskirch et al., 1997; Croxall et al., 1990).

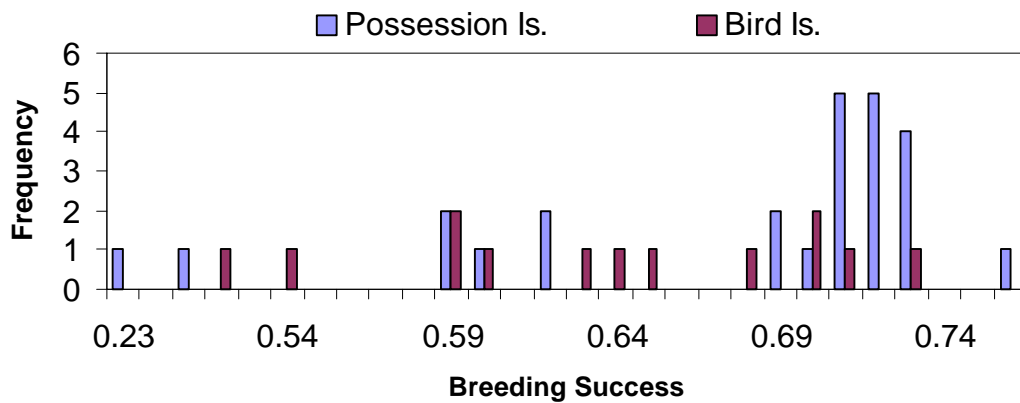


Figure A1-8: Frequency distribution of breeding success on Bird and Possession Islands (data from Weimerskirch et al., 1997; Croxall et al., 1990).



## APPENDIX 2: MODEL

### *Response Variables*

$E_y$  = Eggs produced in year  $y$  per breeding pair

$J_y^{(i)}$  = Juveniles of age  $i$  in year  $y$

$B_y^{(f)}$  = First time breeders in year  $y$

$B_y^{(r)}$  = Repeat breeders in year  $y$

$B_y$  = Total number of breeders in year  $y$

$K_y$  = Non-reproductive adults in year  $y$  (birds skipping reproduction in year  $y$ )

$X_y$  = Successful breeders in year  $y$

$U_y$  = Unsuccessful breeders in year  $y$

### *Driving variables*

$s_y^{juv\ i}$  = Survival rate of juveniles of age  $i$  in year  $y$

$s_y^{ad}$  = Survival rate of adults in year  $y$

$f_y$  = Reproductive success (from egg to fledging) in year  $y$

### *Parameters*

$m_i$  = Probability that a bird of age  $i$  matures and attempts breeding

$fert$  = fertility rate (eggs per nest)

$r_i^{(X)}$  = Proportion of successful breeders that will attempt reproducing next time in year *thisyear+i*

$r_i^{(U)}$  = Proportion of unsuccessful breeders that will attempt reproducing next time in year *thisyear+i*

Eggs per individual breeder or, alternatively, “female” eggs per female

$$(1) \quad E_y = 0.5 \text{ fert } B_y$$

Fledglings (the same year!)

$$(2) \quad F_y = E_y f_y$$

First year juveniles

$$(3) \quad J_y^{(1)} = E_{y-1} s_y^{juv 1}$$

For juvenile birds (those that do not mature)

$$(4) \quad J_y^{(age)} = J_{y-1}^{(age-1)} (1 - m_{age}) s_y^{juv age}$$

First time breeders

$$(5) \quad B_y^{(f)} = \sum_{i=1}^{15} J_{y-1}^{(i-1)} m_i s_y^{juv i-1}$$

Repeat breeders

$$(6) \quad B_y^{(r)} = \sum_{i=1}^4 \left( X_{y-i} r_i^{(X)} \prod_{j=1}^i s_{y-j}^{ad} \right) + \sum_{i=1}^4 \left( U_{y-i} r_i^{(U)} \prod_{j=1}^i s_{y-j}^{ad} \right)$$

Total Breeders

$$(7) \quad B_y = B_y^{(f)} + B_y^{(r)}$$

Successful breeders and unsuccessful breeders

$$(8) \quad X_y = B_y f_y$$

$$(9) \quad U_y = B_y (1 - f_y)$$

Birds skipping reproduction

$$(10) K_y = \sum_{i=1}^4 \left( X_{y-i} \sum_{k=i}^3 \left( r_{k+1}^{(X)} \prod_{j=1}^k s_{y-j}^{ad} \right) + U_{y-i} \sum_{k=i}^3 \left( r_{k+1}^{(U)} \prod_{j=1}^k s_{y-j}^{ad} \right) \right)$$

### *Modeling population natural variability*

Juvenile and adult survival, as well as breeding success are modeled as random variables. Variability in the survival process is modeled as a binomial random process, where:

$$(11) J_y^{(age)} \approx Bin(J_{y-1}^{(age-1)}, surv)$$

This type of error is usually referred to as *demographic stochasticity* in the conservation literature. It represents chance event of individual birds dying when all individuals have the same intrinsic probability. This process represents the minimum possible variability exhibited by population numbers.

Breeding success is modeled by a uniform random variable:

$$(12) f_y \approx Uniform(\min.breed\_rate, \max.breed\_rate)$$

### *Modeling sampling variability*

Sampling variability on the counts of breeders is modeled by adding a log-normal error to the simulated total population size:

$$(13) Obs(B_y) = B_y e^{e+s^2/2}; \quad e \approx Normal(0, s^2)$$

The magnitude of the sampling error is given by  $\sigma$ , which represents the coefficient of variation of the sampling process.

### *Simulating trends in population behavior*

We first tuned the model to produce time series with stable average population numbers (no absolute population change). Additional induced mortality was added to different stages in the population to generate trends of particular sizes (see Appendix 4 for details on parameters used).

## **APPENDIX 3: GUIDE TO SPREADSHEET USE**

This appendix provides a basic guide for the use of the accompanying spreadsheet “Albatross simulation.xls”. This spreadsheet generates 200-year time series of albatross population behavior as a function of different biological parameters and additional induced mortality on juveniles and adults. It also simulates sampling in selected time “windows” and provides estimated trends within this window, as well as test results for the trend with two methods: simple linear regression of  $\ln(\text{numbers})$  on time and maximum likelihood estimation.

### Color codes

Cells are color coded to differentiate initial conditions, basic parameters, derived parameters and calculations.

*Numbers in orange* : population and simulation parameters to be provided by users.

*Text in red*: indicates names assigned to cells which are used to reference them in calculations and input/output operations. These names replace the more traditional reference to cells in spreadsheet operations of the format \$A\$1.

*Numbers in blue*: indicate initial conditions, i.e. initial numbers in different juvenile and adult stages to start the simulations, to be provided by user.

### Layout

Different sections in the spreadsheet are separated by yellow divisions. The main sections are:

*Population parameters*: reproductive and mortality rates defining the basic biology of the species.

Appendix 4 contains a list with the values used for the orange cells in our simulations.

*Simulation parameters*: define additional mortality sources, the period where they occur, whether stochasticity is added or not and to which processes. These are the parameters that define the different scenarios in our simulations. The sampling parameters consist of the type of sampling error (lognormal or binomial), amount of sampling error, the time window where sampling is performed and sampling frequency. Simulation parameters include the number of replicates to be produced, the different window sizes the simulations are to explore, the alpha level for tests, and the name of the ASCII file where simulation results are to be placed. Appendix 4 contains a list with the values used for the orange cells in the different scenarios in our report.

*Dynamics*: this section contains the dynamic simulation of numbers as a function of parameters and initial conditions (marked in blue, to be provided by users).

*Projections*: contains two graphic outputs of one-time projections of population numbers. The first displays breeders and non breeders in the population in the 200 year period. The second displays actual number of breeders in the sampling “window” and the estimated number of breeders after sampling error is included.

### Visual Basic Macros

The dynamics are simulated by the spreadsheet, but Visual Basic macros are used to generate random numbers for vital rates and sampling errors more efficiently and for repetitive calculations in the simulations based on the spreadsheet. For a given set of parameters and initial conditions, macros are run by pressing one of the two gray buttons on the sheet.

### Running modes and outputs

There are two running modes, each associated to one of the gray buttons on the spreadsheet.

#### *One-time projection*

Performs a one-time 200-year projection of the population for given parameters and initial conditions. It allows to visually explore one realization of the simulation process and check results of the tests for that one realization. Changing the sampling window definition determines what is displayed in the graph on the left. It is necessary to run the macro every time any of the simulation parameters is changed.

#### *Simulations*

This mode allows to generate a whole simulation by generating several replicates of the same process. It also allows to generate results for increasing window sizes by changing the value in the cell *intyear*. Increasing window sizes from *intyear* up to *winyears* are tried. For instance, if *fyear* is 100, *winyears* is 40 and *intyear* is 10, the following sampling windows are simulated: 100-110, 100-120, 100-130 and 100-140. If *intyear* contains the same value as *winyears*, a single window of size *winyears* is explored. For instance, if *fyear* is 100, *winyears* is 40 and *intyear* is 40, a single sampling windows is simulated: 100-140.

The results from the simulations are placed in an ASCII file which name is provided by the user in the cell *outfile* (**IMPORTANT**: enter the file name with the complete directory structure for *outfile*

to be created in your working directory, e.g. D:\Albatross\simres.txt) After the simulation is performed, *outfile* will look like:

```

SIMULATION_PARAMETERS
Breeding_success: 0.64
Min.Breeding_success: 0.53
Max.Breeding_success: 0.75
Fertility_rate: 1
rs: 0.1 0.15 0.75 0
ru: 0 0.05 0.15 0.8
Prop_mat_at_age: 0 0 0 0 0 0 0 3.84615384615385E-02 7.47863247863248E-03 0.267094017094017
0.305555555555556 0.191239316239316 7.58547008547008E-02 7.58547008547008E-02 3.84615384615385E-02
Mat_schedule: 0 0 0 0 0 0 0 2.89045812632958E-02 6.2908957572438E-03 0.245757828700351
0.405167635881214 0.463382179505374 0.372298877919194 0.644690090801421 1
Juv.survivals: 0.91 (age_1_to_4 ) 0.92 (age 5 _to_15)
Recruitment_rate: 0.3651829
Ad.survival: 0.93774
Add.ind.juv.and_adult_mortalities: 0.01 0.01 0.01
Years_Incidental_Mort.: 1 200
Stochastic_breed,surv_and_samp.: Y Y Y
Coeff.Var.in_sampling_error: 0.1
Max.window.size_and_samp.frequency: 100 - 139 every 1 years
Test_results_for_windows_every: 5 years
Replicates: 200
Alpha_level: 0.05

SIMULATION_RESULTS
Fields: Replicate Asympt.trend Resid.variance SlopeTests(<>window_sizes) MLETests(<>window_sizes)
1 0.9899375 3.43871E-03 ns ns Y Y Y Y Y Y ns ns ns ns ns ns ns ns
2 0.9901899 3.327467E-03 ns ns Y Y Y Y Y Y ns ns ns ns ns ns ns ns
3 0.9898803 3.536057E-03 ns ns Y Y Y Y Y Y ns ns ns ns ns ns ns ns
4 0.9905379 3.109868E-03 ns ns ns ns ns Y Y Y ns ns ns ns ns ns ns ns
5 0.9899124 3.089036E-03 ns ns ns ns Y Y Y Y ns ns ns ns ns ns ns ns

```

The first 19 lines contain all the population and simulation parameters that define the run. Below come the simulation results which, for each run, consist of a line containing the replicate number, the average change in the breeding population between years 100 and 200, the residual error around the decline in that same period (in log scale) and the results of the tests on estimated breeders. The first block of the test results shows those corresponding to the simple regression approach and for increasing window sizes (in the example above are 5,10,15,...,40), while the second block shows those corresponding to the MLE test. Power curves can be constructed from the proportions of replicates that provide a significant test (“Y”) for each window size. To generate our results, we simulated 500 replicates for each scenario.

## APPENDIX 4: PARAMETERS DEFINING THE SCENARIOS USED IN THE SIMULATIONS.

### *Overall population parameters (invariant across scenarios)*

There is no data available for New Zealand species on juvenile survival, recruitment rates, or variability in numbers over time, but only on breeding success (similar for NZ species and those from South Georgia and Crozet Islands) and adult survival (higher for NZ species, Appendix 1). As shown in this report, the power to detect trends depends primarily on the size of the trend as well as on the amount and structure of variability in numbers between years, more so than on the particular combination of parameters used to generate the trend. We, therefore, preferred to fully parameterize the model based of *D.exulans* data, instead of using a mixture of parameter values from different species and places. This allowed us to better validate the results of our simulations for internal consistency (*i.e.*, we have good estimates for all the parameters of the model and can compare the simulated trends with actual population trajectories).

### Breeding success:

Uniform(0.53, 0.75), average 0.64.

This average value is similar to that estimated for South Georgia and Adams Island. The range is similar to that observed for South Georgia, but was fine tuned to produce a residual variability in counts of breeding pairs comparable to that observed on SG populations (~0,0035 in log scale).

### Proportion of today's breeders that will attempt reproducing next time in $n$ years and maturation schedule

South Georgia data were used to derived these parameters.

Years	Successful breeders	Unsuccessful breeders
1	0	0.8
2	0.75	0.15
3	0.15	0.05
4	0.1	0

Age	Prob. Mature
1	0
2	0
3	0
4	0
5	0
6	0
7	0
8	0.03
9	0.006
10	0.25
11	0.41
12	0.46
13	0.37
14	0.64
15	1

Juveniles survival

Ages 1-4: 0.91, ages 5-15: 0.92

A juvenile survival of 0.91 produces a survival to age 5 of 0.62 (*i.e.*,  $S^5$ ), a value comparable to the newest estimates for Bird and Possession Is. (Fig. A1-4).

Adult survival:

We used a value of 0.9377, which in combination with the chosen maturation schedule, breeding frequencies and juvenile survivals produces a zero population growth rate. This value is comparable to estimates for South Georgia populations, but, as explained before, lower than estimates for NZ species (Appendix 1).



Scenario parameters

SCENARIO	Sampling error CV	Additional mortality adults	Additional mortality young juveniles	Additional mortality old juveniles	Years with additional mortality	Max size for sampling window	Sampling every <i>n</i> years
<b>A. Stabilized trends, sampling every year</b>							
A1. No sampling error. 0% trend 1% trend 2% trend	0	0 0.01 0.0285	0 0.01 0.01	0 0.01 0.01	1-200 1-200	100-139	1
A2 Moderate sampling error. 1% trend 2% trend	0.1	0.01 0.0285	0.01 0.01	0.01 0.01	1-200 1-200	100-139	1
A3 Significant sampling error. 1% trend 2% trend	0.2	0.01 0.0285	0.01 0.01	0.01 0.01	1-200 1-200	100-139	1
<b>B. Stabilized trends, sampling every other year</b>							
B1. Moderate sampling errors. 1% trend 2% trend	0	0.01 0.0285	0.01 0.01	0.01 0.01		100-139	2
<b>C. Transient behavior, sampling every year</b>							
C1. Moderate sampling errors. 1% trend 2% trend	0	0 0	0.0535 0.111	0 0	100-200 100-200	100-139	1