A MICE approach to scoping the possible impact of guano harvests on trends in penguin abundance

D.S. Butterworth and A. Ross-Gillespie [1]

Email: doug.butterworth@uct.ac.za

Summary

A simple population dynamics model is used the scope the possible impact of guano harvesting during the 20th century on penguin population trends, as a consequence of the associated reduction in optimal breeding habitat diminishing reproductive success. The analyses' outputs include the delays arising from transient effects. They indicate that guano harvesting (which ended in 1992) could have been a major contributing factor to the reduction in penguin abundance over the last few decades of the 20th century. However, transient effects become minimal after the turn of the that century; such effects could not (even partially) account for the decline in penguin counts after about 2015.

Keywords: population decline impact, guano harvest, MICE, penguin

Introduction

A number of hypotheses have been advanced to explain (at least in part) the generally continuing decline in penguin numbers in South Africa since the middle of the last century (e.g. Butterworth and Ross-Gillespie, 2022). One of these is that guano harvests from islands during the last century decreased the optimal breeding habitat for penguins, leading effectively to a diminishing "carrying capacity" for penguins.

A major immediate conservation concern is that this penguin decline continues to the current time (the 2020's). Since guano harvesting ended in 1992, it might be thought that this cannot be a contributor to penguin declines over the last three decades. However, that ignores the fact that since the guano harvests impact *reproductive success* through loss of better breeding habitat, the impact on the mature population will be delayed as the new chicks first need to age before they join the mature component of the population. This will lead to transient effects with the total population size continuing to drop for some time after the guano harvests ceased.

The purpose of this paper is to use a simple MICE approach to investigate this matter quantitatively in two respects: the extent to which guano harvests might explain the decline in the penguin counts historically, and the role which transient effects might play in this in more recent years. The approach is deliberately broad brush and coarse for what is an initial scoping exercise. Thus, for example, spatial variation in both penguin counts and guano harvests are ignored for this broad brush level analysis, which aggregates over all penguin colonies and all guano harvests.

Data

The data that have been used in these analyses are the penguin count data (number of breeding pairs for 19 colonies, for the years 1956-2021, but with many gaps during the years 1957-1990), and the annual guano

[1] Marine Resource Assessment and Management Group, Department of Mathematics and Applied Mathematics, University of Cape Town, Rondebosch 7701.

harvest from 10 islands from the years 1896-2000. These data sets have kindly been provided by A. Makhado of DFFE (*pers. comm.*). Figure 1 plots the cumulative guano harvest.

Methods and Results

Appendix A sets out the algebraic specifications of the MICE used, and Appendix B details how gaps in the time series of penguin counts are filled to allow a time series for the total penguin population size to be obtained.

Clearly, notable effects will be evident only if the guano harvest has had a substantial impact on the penguins' optimal breeding habitat, so that the results shown are restricted to a range of lowish values of the parameter γ, which is the proportion to which the penguin carrying capacity has been reduced by the time harvesting ceased in 1992 (see Figure 1). The range of 0.025 to 0.2 is considered, with 0.05 serving for a base case.

The base case choices for the other two key parameters are first survival proportion S = 0.85 – this is motivated as a rough average of the estimates for Robben Island penguins in the 1990s (Robinson *et al*, 2015). For H, a seemingly relatively high value of 0.7 is chosen for the base case, but bear in mind that reproduction is assumed to be density dependent, so that this is the value that would apply in the extreme of penguin abundance very far below its carrying capacity.

Results have then been developed around single factor variations about the base case of γ=0.05, S = 0.85, and H = 0.7.Table 1 lists estimates of average rates of decline for the observed and model-predicted penguin population sizes for a selection of time periods for these different parameter values choices around the base case. Figure 2 plots the corresponding observed and model predicted penguin population trajectories.

Discussion

A few general observations suggest themselves following inspection of the results in Table 1 and Figure 2.

- Interpretation of how well different parameter options for the model fit the data is rendered difficult by the relatively large peak in penguin abundance around the turn of the 20th century, which clearly must be a result of other mechanisms (with the upward pulses in both sardine and anchovy abundance at that time almost certainly playing some role).
- Modelled penguin population trends are most sensitive to the value specified for γ (the proportion to which guano harvesting has reduced penguin carrying capacity through limiting optimal breeding habitat). Sensitivity is less to the value selected for S, and less still to that for H.
- The value of γ needs to be about 0.05 or less to see modelled population trends approaching the magnitude of the declines indicated by the more recent penguin count data.
- Improvements in the fits to the count data when varying the values of the model parameters relate primarily to the 1980-2000 period.
- For most of the scenarios considered, any transient effect from the guano harvesting is small after the turn of the 20th century.
- Penguin count and population model trends are both fairly flat over the 2010 to 2015 period, but these guano-linked models cannot reproduce the declining trend shown by the count data thereafter.

Further work on this topic might include sensitivities to possible levels of guano harvest before the data series starts in 1896, and alternative approaches to estimating individual penguin colony sizes in the years between census counts (i.e., alternatives to the method outlined in Appendix B). However, there are no penguin abundance estimates that far in the past, and these sensitivities would hardly impact the recent trends that are the primary focus of what is intended as a scoping (strategic) exercise only.

Further wider work on this topic will, in the first instance, consider the possible impacts of forage fish abundance and seal predation on penguin trends, with such models being fitted to the penguin trend data. That work may then be extended to models incorporating multiple causes, at which time it would probably be most appropriate to return to considering this guano breeding habitat model further.

Conclusion

These scoping analyses indicate that guano harvesting could have been a major contributing factor to the reduction in penguin abundance over the last few decades of the 20th century. However, transient effects become minimal after the turn of the that century, and in particular such effects could not (even partially) account for the decline in penguin counts after about 2015.

References

Butterworth, D.S. and Ross-Gillespie, A. 2022. Comment on “South Africa’s experimental fisheries closures and recovery of the endangered African penguin” by Sydeman et al. (2021). ICES Journal of Marine Science, 2022, 79, 1965-1971. https://doi.org/10.1093/icesjms/fsac113.

Robinson, W. M. L., Butterworth, D. S. and Plagányi, É.E. 2015. Quantifying the projected impact of the South African sardine fishery on the Robben Island penguin colony. ICES Journal of Marine Science, 72: 1822–1833. https://academic.oup.com/icesjms/article/72/6/1822/919990.

Table 1: Average annual penguin rates of decline calculated as $(1/n) * ln\,[N(y)/N(y-n)]$, where *N* is the observed or model estimated penguin population size and *n* is the number of years between the start and end year of the period considered. Results are shown for the penguin data (bold text) where *N* is the penguin count, as well as for the model estimated population size $\widehat{N}$. There are three sections: (a) *S* (0.85) and *H* (0.7) are fixed and γ is varied, (b) *H* (0.7) and γ (0.05) are fixed and S is varied and (c) *S* (0.85) and γ (0.05) are fixed and *H* is varied. These results are reported for three periods: all the years for which penguin count data are available (1956-2021), and two reduced periods covering recent years of (1980-2021) and (1980-2015).

	Vary γ	All years	1980-2021	1980-2015		Vary S	All years	1980-2021	1980-2015		Vary H	All years	1980-2021	1980-2015
	Data	-3.49	-4.00	-2.87		Data	-3.49	-4.00	-2.87		Data	-3.49	-4.00	-2.87
Model	γ=0.025	-3.68	-2.81	-3.29	Model	S=0.8	-3.00	-2.41	-2.76	Model	H=0.5	-2.92	-2.19	-2.54
	γ=0.050	-2.73	-1.79	-2.10		S=0.85	-2.73	-1.79	-2.10		H=0.7	-2.73	-1.79	-2.10
	γ=0.100	-1.86	-1.04	-1.22		S=0.90	-2.59	-1.44	-1.69		H=0.9	-2.63	-1.58	-1.85
	γ=0.200	-1.11	-0.53	-0.62										

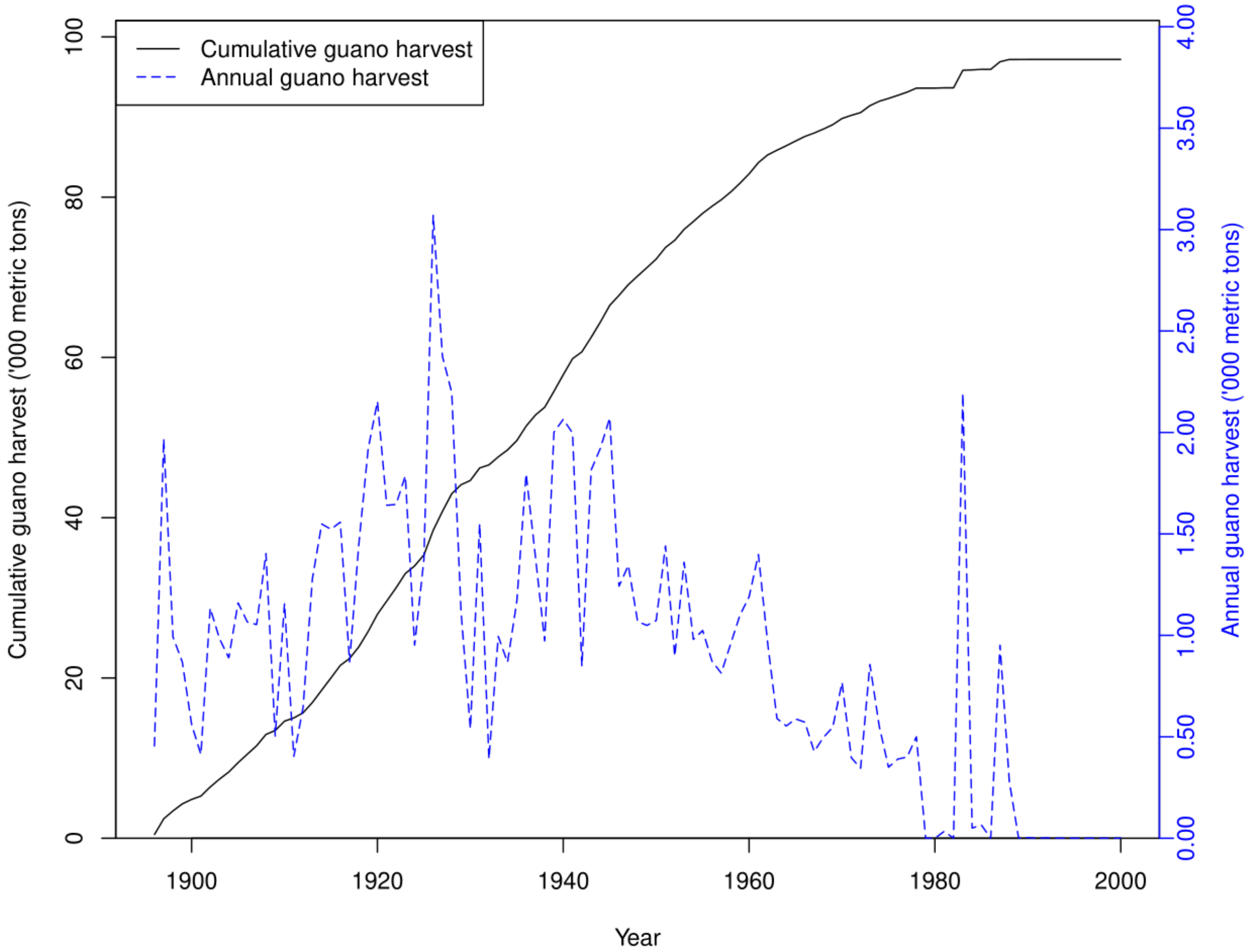

Figure 1: Plot of the guano harvest in thousand metric tons, as kindly provided by A. Makhado, DFFE (*pers. comm.*). The black line shows the harvest considered cumulatively (i.e., total combined harvest up to and including the year indicated). The actual annual harvest is shown by the blue dashed line.

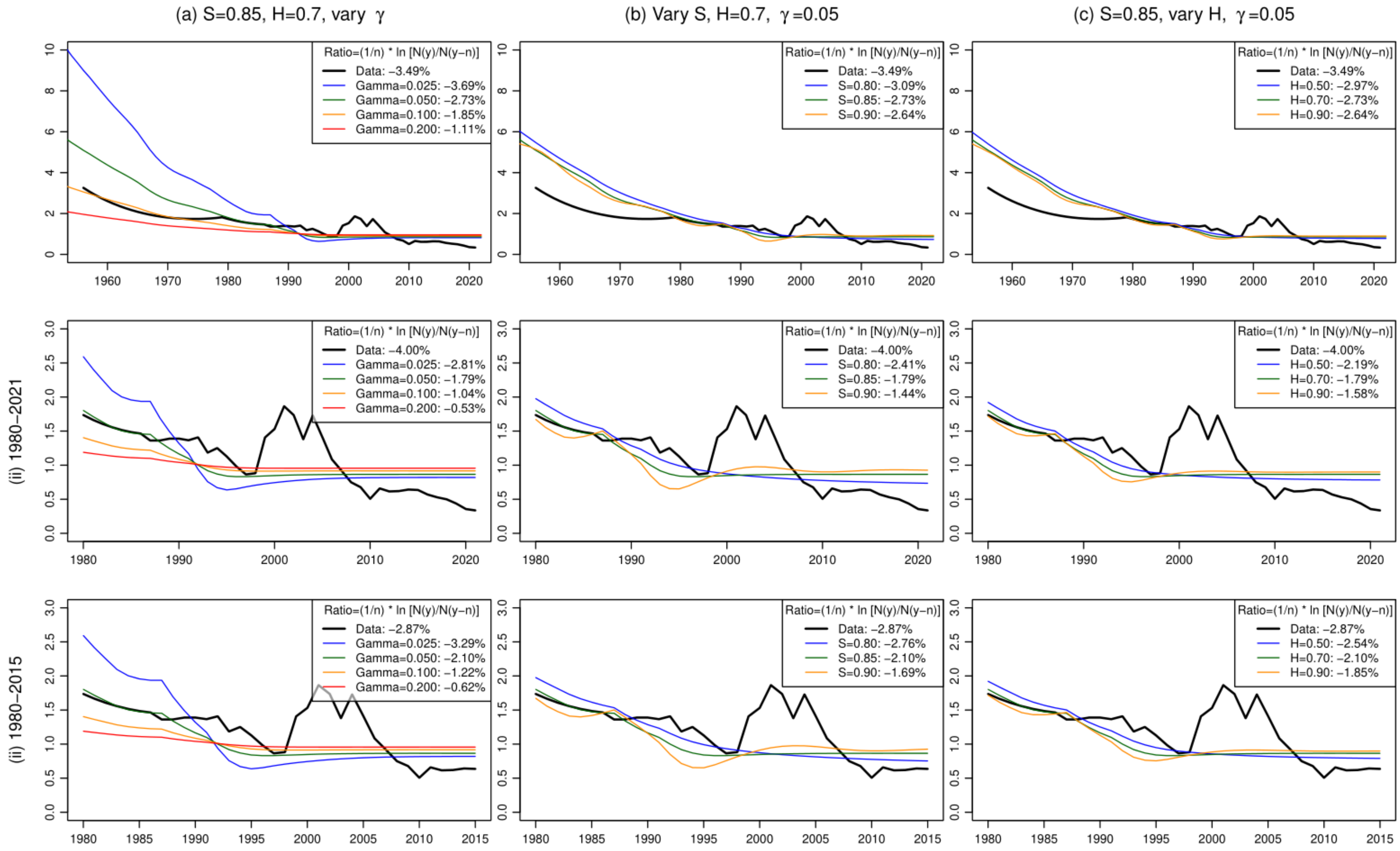

Figure 2: Figure giving population trajectories corresponding to Table 1, showing results for a range of H, S, γ values, and for a selection of three different time periods. The count data and the model population estimates have both been normalised to have the same geometric mean over the period 1980-2021.

Appendix A: Penguin-guano MICE equations

The number of mature penguin pairs in year *y*, N_y, is given by:

$$N_{y+1} = N_y S + N_{y-3} \acute{H} \left(1 - \frac{N_{y-3}}{b_{y-3}}\right) S_j S^3 \quad \text{(A1)}$$

where

- S is the post first year annual survival proportion,
- S_j is the juvenile (first year) survival proportion,
- $\acute{H}$ is essentially the number of fledged chicks per pair, and
- b_y is a measure of carrying capacity which could be reduced, for example, by a decreased optimal breeding habitat as a result of guano harvesting.

Note that an age-at-maturity of three years has been assumed here. Also note that $\acute{H}$ can at maximum be one (i.e., two chicks – "one further pair" raised successfully to fledging age per pair each year), though in reality this limit would not be reached.

Define $H = \acute{H} S_j / S$. Equation (A1) can then be re-written as:

$$N_{y+1} = N_y S + N_{y-3} H \left(1 - \frac{N_{y-3}}{b_{y-3}}\right) S^4 \quad \text{(A2)}$$

Hence H is the number of fledged chicks per pair, adjusted by the ratio of their survival proportion for the rest of their first year of life to the adult survival proportion. This adjustment allows calculations to be carried out without the complication of having to make separate assumptions about the juvenile survival proportion, and effectively assumes that if a parent dies, the chick dies too.

The time-varying carrying capacity, b_y, is assumed to be linearly related to the amount of guano present on the islands (taken to be negatively related to the cumulative guano harvest):

$$b_y = \alpha + \beta G_y \quad \text{(A3)}$$

where:

- α, β are parameters[2], and
- G_y is a measure of the amount of guano present in year *y*::

$$G_{y+1} = G_y - G_y^{harvest} \quad \text{(A4)}$$

where $G_y^{harvest}$ is the amount of guano harvested in year y, and G_0 is the amount of guano present before guano harvesting commenced. To calculate G_0, we assume $G_y = 0$ in the last year during which guano was harvested, and work backwards to add the amount of guano harvested each year to obtain an estimate of G_0. This is approximate because there would still be some new guano added each year, but this amount is probably small by comparison.

We define γ as the proportion to which the penguin carrying capacity is reduced once guano harvesting ceased, i.e.:

$$\gamma = \alpha / (\alpha + \beta G_0) \quad \text{(A5)}$$

Hence α and β can be redefined as:

$$\alpha = \gamma b_0 \quad \text{(A6)}$$

$$\beta = b_0 (1 - \gamma) / G_0 \quad \text{(A7)}$$

[2] Note that α, and hence γ, subsume the effect of the amount of any guano remaining after harvesting ceased.

Furthermore, b_0 can be calculated by assuming initial equilibrium before guano harvesting commences in Equation (A2):

$$b_0 = N_0/[1 - (1 - S)/(HS^4)] \quad \text{(A8)}$$

Hence the model results will depend on the values of only four parameters: N_0 , H, S and γ.

For simplicity a N_0 has been set to 1 so that it is the relative decline in the penguin population which is modelled. Furthermore, the model is started in 1890 (first guano harvest is recorded in 1896) so that the assumption that $N_3 = N_2 = N_1 = N_0$ and similarly that $b_3 = b_2 = b_1 = b_0$ can be made.

Appendix B: Method for filling gaps in the penguin count data set

Penguin population data are available from 1956-2021 for 19 different colonies, as provided by A. Makhado (*pers. comm.*). Figure B1 plots these data. In order to obtain the total population counts used to (potentially) fit annual data to models and to calculate the rates of decline reported for the data in Table 1 of the main text, the gaps in the data were filled by assuming an exponential trend between the nearest years with data. For example, given a series of population counts N_i for year y_i, if two data points are (y_1, N_1) and (y_2, N_2) with missing years between $y1$ and $y2$, then these missing values are determined by the formula:

$$lnN_i = m * y_i + c \tag{B1}$$

where $m = (lnN_2 - lnN_1)/(y_2 - y_1)$ and $c = lnN_2 - m * y_2$.

The resulting curves have been superimposed as red curves onto the plots of the data in Figure A1. For the analyses in this paper, a measure of the total population count has been used, which was calculated by summing the individual gap-filled curves over all the colonies for each year.

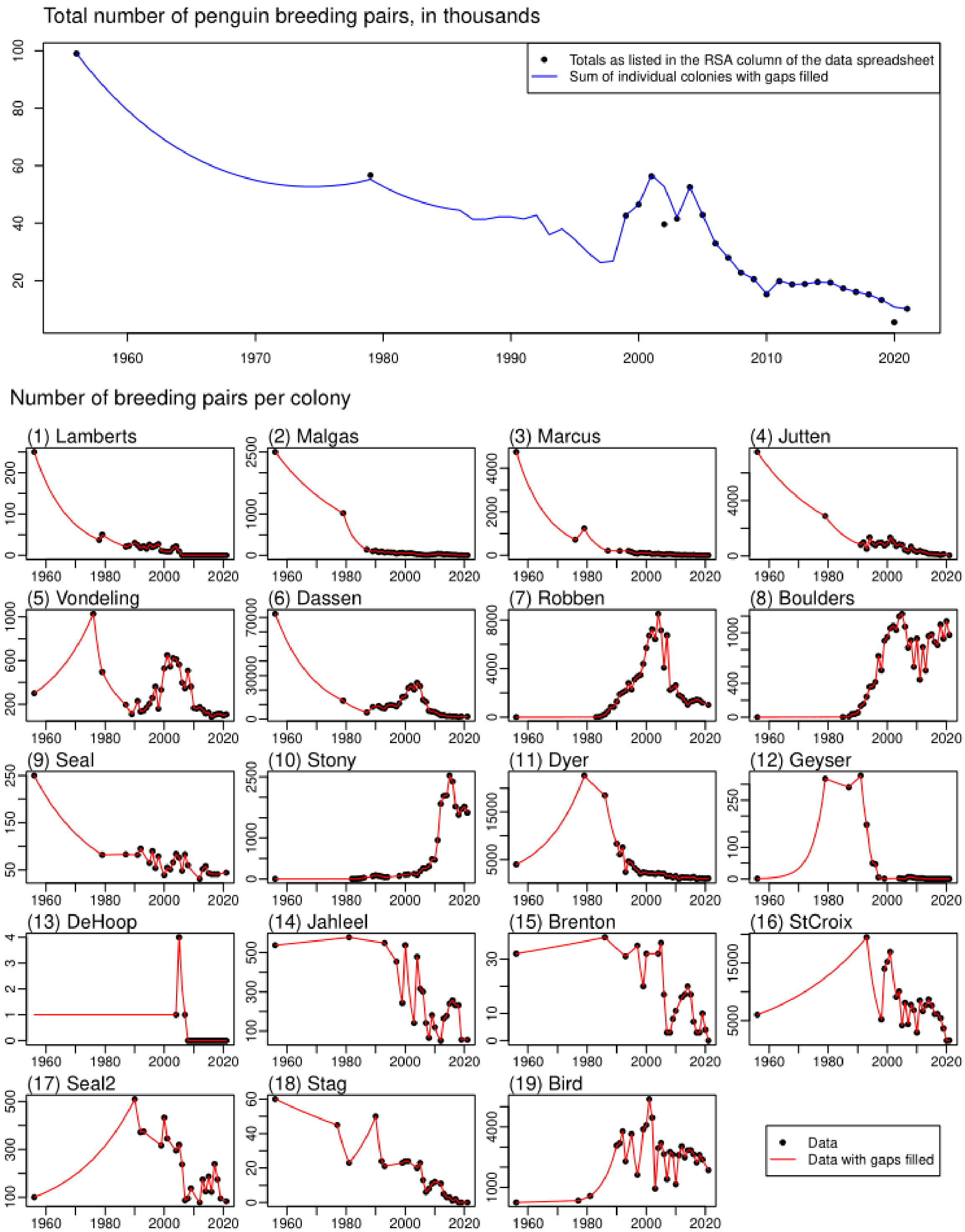

Figure B1: Plots of the penguin count data (number of breeding pairs) as provided in the spreadsheet "African penguin Population data 2022.xlsx" kindly provided by A. Makhado (*pers. comm.*). Note that the top plot corresponds to a column in this spreadsheet marked "RSA". For all years other than 1979, the RSA column corresponds to the sum of the available data across the individual colonies. For certain years where there are more missing data points (e.g., the years 2002 and 2020) the value in the RSA column is lower than the sum of the individual colonies with gaps filled, as the RSA column essentially ignores the missing colonies while the latter attempts to incorporate them by filling the gaps. For the individual colony plots, the curves resulting when the gaps are filled are shown in red. These individual curves have been summed to obtain the blue curve in the top plot, which has been used for the analyses of this paper.