

Re-analysis of the island closure experiment results to implement the suggestions of the December 2020 International Panel

A. Ross-Gillespie and D.S. Butterworth¹

Contact email: andrea.ross-gillespie@uct.ac.za

Summary

The suggestions of the December 2020 International Panel for further analysis of the results from the island closure experiment, in particular to use the same data within a common framework to facilitate comparisons, and to include month as a covariate, are implemented for all response variables except chick survival. Broadly, the results indicate that aggregated and disaggregated approaches give the same results for estimates and variances of the island closure effect (see Figure 2). This is especially a consequence of implementing the disaggregated approach with a nesting structure as advocated by the Panel, which is shown to be statistically justified. Inclusion of the month co-variate does impact results to some extent, the more so for foraging data for the west coast islands (see Figure 3). The sensitivities investigated generally make little difference to results; these checks include extensions to incorporate data prior to 2008. As to be expected, REML based estimates of CIs are somewhat wider than those based on MLE.

Key words: penguin, island closure, fishing impact, month, GLM standardisation

Introduction

This document provides results for implementing the suggestions of the December 2020 International Panel for further analysis of the results from the island closure experiment, in particular to use the same data within a common framework to facilitate comparisons. It includes all response variables except one (chick survival which will be addressed in a separate document by Bergh and colleagues). Specific aims of this document are three-fold: (1) to compare results for annually aggregated and disaggregated (individual) data; (2) to standardise for month where this information is available and assess the impact of this standardisation on the estimates of the fishery-effected change in penguin population growth rate; and (3) to evaluate the impact of the pre-2008 data as well as some other sensitivities on the estimates of change in population growth rate.

Methods

Appendix A contains extracts from Ross-Gillespie and Butterworth (2021); these include a more detailed overview of the general methods applied in this document.

Month information has been provided for chick condition, chick growth, maximum foraging distant for the West Coast (WC) and East Coast (EC) (except for chick growth data which is not available for the EC). There are many ways in which standardisation for month could be implemented. The one adopted here was chosen to enable comparable application to the aggregated and disaggregated approaches. To effect this standardisation, the following GLM was applied to the disaggregated data:

$$f(F_{i,y,m}) = k_i + \alpha_{i,y} + \beta_{i,m} \quad (1)$$

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

where F is the response variable, k is a constant, α_y the year effect and β_m the month effect. Chick condition and chick growth are analysed in normal space (i.e. $f(F) = F$) as both datasets contain non-positives. The foraging data are analysed in log space (i.e. $f(F) = \ln F$). The GLM is applied to each island separately. The standardised response variable is given by:

$$\bar{F}_{i,y,m} = F_{i,y,m} - \beta_{i,m} \quad (2)$$

The island-closure GLM analysis is then applied to the standardised disaggregated data, as well as to standardised aggregated data which are derived by aggregating (annually averaging) the standardised disaggregated data. Additionally, outliers (identified by residuals lying outside the $\mu \pm 3sd$ range from the standardisation in Equation (1)) are removed. Two methods for removal were considered: (M1) outliers are identified for each island/year combination and (M2) outliers are identified across all years for each island. Given the relatively small sample sizes for some of the response variables, it seemed more defensible to calculate outliers over the whole data set rather than splitting it further by year to calculate outliers; consequently method (M2) was preferred.

Estimates and se's of the fishing effect δ are converted to estimates of change in population growth rate (expressed as an annual proportion) by multiplying the δ estimates and se's by 0.094 (for all variables other than chick condition) and 0.11 for chick condition (see Appendix B). Note that for chick condition there is a basis (inferred from results for the macaroni penguin population) to relate the response variable to population growth rate. For the other variables considered, except for fledging success for which a direct relationship follows from the equation for the penguin population dynamics), an assumption is made of a linear relationship through the origin.

Note that for chick growth some 2% of the total observations are non-positive. As this variable is assessed in the island-closure GLM in log-space, the lowest 12% of the values for each island/year combination were removed after the standardisation and before applying the island-closure GLM; this is to retain balance in the manner in which these data are censored to reduce possible bias in the resultant estimates of the island closure effect (see

Table 1).

Basic results are shown for seven estimation approaches, plus an additional eighth approach for chick growth; sensitivities are also reported for some of these.

	Approach	Underlying data
A1	Standard island closure GLM (see Equation 1 of Appendix A) with response variable analysed in log space for all variables except chick condition, which is analysed in normal space	Original aggregated (i.e. non-standardised) data set. For chick growth, the lowest 12% have been removed for each island/year combination, thus excluding negative values in a balanced manner.
A2	Same island-closure GLM as for A1	Standardised aggregated data, with outliers removed according to method M2. For chick growth, the lowest 12% have been removed for each island/year combination.
A3	Approach A1 but with sample-size-weighted and island-dependent variance (see Appendix A)	Standardised aggregated data, with outliers removed according to method M2. For chick growth, the bottom 12% have been removed for each island/year combination, i.e. same as for A2
A1*	Same island-closure GLM as A1	Applicable to chick growth only – replicate of A1, i.e. original data but with the lowest 12% not removed before the data are aggregated. After aggregation, all inputs to the island closure GLM are positive
D0	Unnested model, with random effects components of Island:Year for chick condition and growth and Island:Year:BirdID for the foraging data.	Original disaggregated data. Lowest 12% per island/year combination removed for chick growth.
D1	Nested model with a (Year/Island) random effects component.	Original disaggregated data. Lowest 12% per island/year combination removed for chick growth.
D2	Nested model with a (Year/Island) random effects component.	Standardised disaggregated data with outliers removed according to method M2. Lowest 12% per island/year combination removed for chick growth.
D3	Model D2 but with island-dependent variance	Standardised disaggregated data with outliers removed according to method M2. Lowest 12% per island/year combination removed for chick growth.

Results

Table 2 lists the estimates of the change in population growth rate (expressed as an annual proportion) associated with fishing compared to island closure (note therefore that this is a “fishing effect”, i.e. if an island being open to fishing has a negative impact on the penguin population, this value will be negative) for the aggregated approach A3 and the disaggregated approach D3. Both approaches A3 and D3 have island-dependent variance and use standardised data, so this Table aims primarily to compare the use of aggregated vs disaggregated data approaches as best corresponding to the basis suggested by the December 2020 International Panel.

Figure 1 shows the Zeh plots for all eight approaches, and Figure 2 provides the same A3 vs D3 comparison as given in Table 2. Table 3 lists the estimates for approaches A1 and A2. As both approaches use the same GLM but A1 uses non-standardised and A2 standardised data, the primary purpose of this Table is to show the impact of the standardisation for month. Figure 3 shows the corresponding Zeh plots. Table 4a and Table 4b list the full set of results for the estimated change in population growth rate for all the approaches and all the response variables for which disaggregated data are available. Figure 1a and Figure 1b show the corresponding Zeh plots.

As no month data were available for fledging success, month-standardisation could not be conducted for this response variable. So in order to include the fledging success results for comparison purposes, a variant of the A3 approach, which implements sample-size-weighted and island-dependent variance but uses non-standardised aggregated data, was applied to all the response variables including fledging success. Additionally, where data are available prior to 2008 (in the interest of consistency, all other results in this document correspond to data from 2008 onwards only), the modified A3 approach was applied to the aggregated data for the whole period. Table 5 lists these results and Figure 4 shows the corresponding Zeh plots.

A final sensitivity run was to remove data entries from the foraging datasets corresponding to multiple trips taken by the same bird, i.e. if a particular bird took more than one foraging trip, only the first record was retained. The inclusion of these multiple trips strictly speaking lead to non-independent data points, which could lead to negatively biased CI intervals. The sensitivity was conducted to evaluate the impact of removing the multiple trips. Table 6 lists the results.

Table 7a compares results for approaches A2 and D3 using the maximum likelihood (MLE) method for fitting the GLM with the restricted maximum likelihood (REML) method. The MLE method is the standard approach for the results presented in this document, as the A3 approach is coded in ADMB for which REML estimation is not available. Table 7b lists the differences in the se's estimated by the MLE and REML methods.

Discussion

Key aspects of the results are as follows. Note that the authors have attempted to keep the points made below "opinion-free". Further views with arguably some subjective element are provided in Butterworth and Ross-Gillespie (2021).

- Use of aggregated vs disaggregated data:
 - In Figure 1, the A1 and D0 approach results are representative of (although for some variables not identical to) the understanding in December 2020 for "best" implementations of the aggregated and disaggregated approaches respectively. Since that time, however, there have been a few changes in the way in which the GLM-estimated fishing effect parameter δ is reported. These δ parameters are now converted into estimates of the annual change in population growth rate (see Appendices A and B), as this is a more readily interpreted quantity in terms of its implications, and it can be reported in a comparable manner across the different response variables. Additionally, a relationship between chick survival and chick condition has been developed (see Appendix B) which allows chick condition and the change in population growth rate to be linked. Apart from the desirability of being able to make use of such a relationship, this also became necessary because the disaggregated data for chick condition contain negative values, and this direct relationship allowed for the disaggregated data to be analysed in normal space, avoiding the problem of being unable to readily incorporate negative values in log-space.
 - The A3 and D3 approaches are put forward as representing the best implementations of the suggestions of the December 2020 International Panel.
 - The point estimates for change in population growth rates are very similar for the aggregated (A3) and disaggregated (D3) approaches. The estimates for either approach are well within the CI for the other (see Table 1 and Figure 2). This broad result is in line with the expectation expressed by the December 2020 International Panel.
 - For chick condition and chick growth, the CIs are also similar, though the CIs estimated for the foraging data tend to be 20-40% smaller for the foraging data when the disaggregated data are used. This similarity is also generally in line with Panel expectations. In principle the same results are to be expected from both approaches (Butterworth 2020); the differences seem associated especially with instances of small sample sizes in some years for the foraging data. Note that the differences are much less than appeared to be the case in December 2020 (here the D0 approach is considered to be representative of the disaggregated approach for which results were presented in December).
 - Some negative aspects of both (aggregated and disaggregated) approaches are:
 - Aggregated – The ample size adjustment is approximate, and likely does not work as well at very low sample sizes.
 - Disaggregated – This approach may not allow fully for pseudo-replication, which would result in smaller CIs than are appropriate. Estimates of precision can be sensitive to inappropriate selections for hierarchical structure (nesting).
 - Both – The use of random effects (also included for year in aggregated approaches to robustify estimation given limited number of degrees of freedom) can be problematic

because there are indications of estimation instability as the data are not well able to partition the contributions to the variance of different effects (see for example the considerable imprecision indicated for the estimates of the variances of the various random effects which are reported in Table 6a of Appendix A).

- Month standardisation
 - The month standardisation makes very little difference for chick condition and growth.
 - There is some impact for foraging data – CIs can decrease up to 60% for WC islands (e.g. Path length for Dassen).
 - The largest decreases in CI's are, however, a consequence of removing outliers (compare the results for approaches A1, A1.5 and A2 in Table 3).
- Sensitivities
 - Incorporation of pre-2008 data
 - This does not make a substantial impact, with the notable exception that the estimates of the island closure effect switch signs for both Dassen and Robben island fledging success when the pre-2008 data are included (see Table 5 and Figure 4).
 - In principle, one would expect that a longer series gives more degrees of freedom and hence that better precision could be expected, but this not reflected in the results. There may be some confounding arising with all pre-2008 situations regarded as the neighbourhood of the island being open to fishing, but some catches during this period being rather small. The results in this document are all based on a "closure" estimator rather than a "catch" estimator (which would have taken that catch size into account); the reason for the former to be preferred is that the latter did not perform as well in simulation tests because of the possible positive correlation between the size of the catch and the biomass of forage fish from year to year (Ross-Gillespie and Butterworth 2016).
 - Multiple trips by the same bird
 - Table 6 lists the results for when only the first trip undertaken by any bird is retained in the data set, and future (repeat) trips by that same bird that year are ignored in order to offset the effect from non-independence in the data set. The impact of this removal is minimal.
 - REML vs MLE
 - For most cases, the actual point estimates themselves are very similar for the REML and the MLE methods. For the aggregated data approach A2 they are virtually identical, while for the disaggregated approach D3 the differences are small, with the largest difference being for Dassen trip duration where the REML point estimate of change in population growth rate is more positive (2.5% for REML, 3.8% for MLE) (see Table 7a).
 - In all cases, the estimate of precision is larger for the REML method, as would be expected. The $se(REML)$ is 6% - 77% larger than the $se(MLE)$, with the largest of these differences occurring for the disaggregated approach (see Table 7b).
 - REML is the preferred method as it provides unbiased estimates of variance, but since REML is not available in ADMB, which is needed to code for the sample-size-dependent variance adjustment, the MLE method has been used for most of the results presented in this document.

References

- Butterworth, D.S. 2020. A response to Sherley: FISHERIES/2020/JUL/SWG-PEL/53REV. DEFF Fisheries document FISHERIES/2020/AUG/SWG-PEL/82.
- Butterworth, D.S. and Ross-Gillespie, A.. 2021. A summary of results for the island closure experiment. DEFF Fisheries document FISHERIES/2021/JUN/SWG-PEL/36.
- Coetzee, J. and Merkle, D. 2020. The current set of available data for evaluation of the Island Closure Experiment. DEFF Fisheries document FISHERIES/2020/SEP/SWG-PEL/100.
- Ross-Gillespie, A. and Butterworth, D.S. 2016. Implementation of the Algorithm recommended by the Panel for the 2016 International Stock Assessment Workshop for assessing whether or not to continue with the penguin island closure experiment. DAFF FISHERIES document FISHERIES/2016/DEC/SWG-PEL/77rev.
- Ross-Gillespie, A. and Butterworth, D.S. 2021. An initial implementation of suggestions by the 2020 Panel to improve estimates of the effect of fishing around islands on penguins by using models with random effects and applied to both aggregated and disaggregated data. DEFF Fisheries document FISHERIES/2021/APR/SWG-PEL/24.

Table 1: Break down of the number of non-positive observations by island and year. Although a total of only about 2% of the total observations are non-positive, just over 13% of the Dassen 2014 observations are non-positive. In order therefore to remove non-positives consistently by island and year (to avoid introducing possible other sources of bias) by deleting the lowest X% for each island and year pair, the smallest X possible is 12% (the number removed is rounded, so $30 \times 12\% = 3.6$, which has been rounded to 4).

Island	Year	Total N	N<=0	N<=0/N*100
Dassen	2008	14	0	0
	2009	70	1	1.43
	2010	45	0	0
	2011	19	1	5.26
	2012	126	0	0
	2013	70	7	10
	2014	30	4	13.33
Robben	2008	50	1	2
	2009	88	0	0
	2011	54	0	0
	2012	80	0	0
	2013	75	0	0
	2014	39	1	2.56
Total		760	15	1.97

Table 2: Estimates of change in population growth rate (expressed as an annual proportion) for approaches A3 and D3 (i.e. comparison of results obtained by applying the island-closure GLM to aggregated and to disaggregated data). Note that for some response variables, the estimated random effects variance parameter is rather small, in particular for EC trip duration, where $\sigma_{\alpha} = 0.0002$ for approach A3. The analysis for this case was re-run setting a lower bound of 0.05 on σ_{α} ; this had minimal impact on the estimated change in population growth rates, which changed from -0.145 (se 0.0072) to -0.0143 (se 0.0073) for Bird and from 0.0003 (se 0.0071) to -0.0005 (se 0.0072) for St Croix, so for consistency results without this bound are reported below.

		A3				D3				A3-D3				Average no. data points per year
		Estimate	L95	U95	Range (A3) Upper-Lower	Estimate	L95	U95	Range (D3) Upper-Lower	Mean(A3)- Mean(D3)	Range(A3)- Range(D3)	(Mean(A3)- Mean(D3)) /Range(D3)	(Range(A3)- Range(D3)) /Range(D3)	
Dassen	Condition	0.001	-0.007	0.008	0.015	0.000	-0.008	0.007	0.015	0.001	0.000	0.08	0.02	474
Dassen	Growth	0.019	-0.006	0.043	0.049	0.020	-0.007	0.047	0.054	-0.001	-0.005	-0.03	-0.10	47
Dassen	MaxDist	-0.008	-0.029	0.013	0.042	-0.010	-0.026	0.005	0.031	0.002	0.011	0.08	0.34	21
Dassen	PathLength	0.013	-0.006	0.033	0.039	0.005	-0.011	0.021	0.032	0.008	0.008	0.26	0.24	21
Dassen	TripDuration	0.031	0.004	0.057	0.053	0.025	0.004	0.046	0.042	0.006	0.012	0.13	0.28	21
Robben	Condition	-0.009	-0.017	-0.002	0.015	-0.010	-0.018	-0.003	0.015	0.001	0.000	0.04	0.02	527
Robben	Growth	0.003	-0.023	0.029	0.052	0.001	-0.026	0.028	0.053	0.002	-0.001	0.03	-0.02	56
Robben	MaxDist	-0.002	-0.022	0.018	0.040	-0.002	-0.017	0.012	0.029	0.001	0.011	0.02	0.38	27
Robben	PathLength	-0.008	-0.026	0.010	0.036	-0.009	-0.023	0.004	0.027	0.001	0.009	0.04	0.35	25
Robben	TripDuration	0.004	-0.022	0.030	0.052	0.001	-0.019	0.021	0.040	0.003	0.012	0.07	0.31	26
Bird	Condition	-0.003	-0.017	0.012	0.029	-0.004	-0.019	0.011	0.030	0.002	-0.001	0.05	-0.02	207
Bird	MaxDist	0.002	-0.010	0.014	0.024	0.007	-0.005	0.019	0.024	-0.005	0.000	-0.21	0.00	57
Bird	PathLength	0.009	-0.007	0.025	0.032	0.005	-0.013	0.024	0.037	0.004	-0.005	0.11	-0.14	45
Bird	TripDuration	-0.014	-0.029	0.000	0.029	-0.006	-0.023	0.011	0.034	-0.008	-0.006	-0.25	-0.17	53
StCroix	Condition	0.007	-0.007	0.020	0.027	0.003	-0.012	0.018	0.029	0.004	-0.002	0.12	-0.06	135
StCroix	MaxDist	-0.034	-0.047	-0.021	0.026	-0.04	-0.054	-0.027	0.027	0.006	-0.001	0.24	-0.03	22
StCroix	PathLength	-0.021	-0.037	-0.004	0.033	-0.026	-0.044	-0.007	0.037	0.005	-0.004	0.14	-0.11	17
StCroix	TripDuration	0.000	-0.014	0.014	0.028	-0.005	-0.021	0.011	0.033	0.005	-0.004	0.17	-0.14	19

Table 3: Estimates of change in population growth rate (expressed as an annual proportion) for approaches A1 and A2 (to show the impact of standardising for month).

		A1				A2				A1-A2				Average no. data points per year
		Estimate	L95	U95	Range (A1) Upper-Lower	Estimate	L95	U95	Range (A2) Upper-Lower	Mean(A1)- Mean(A2)	Range(A1)- Range(A2)	(Mean(A1)- Mean(A2))/ Range(A2)	(Range(A1)- Range(A2))/ Range(A2)	
Dassen	Condition	-0.002	-0.010	0.006	0.015	0.001	-0.007	0.008	0.015	-0.003	0.000	-0.19	0.02	474
Dassen	Growth	0.018	-0.002	0.039	0.042	0.015	-0.006	0.035	0.041	0.004	0.001	0.09	0.02	47
Dassen	MaxDist	0.004	-0.026	0.035	0.061	-0.001	-0.021	0.019	0.040	0.005	0.021	0.14	0.53	21
Dassen	PathLength	0.022	-0.007	0.050	0.057	0.019	0.001	0.037	0.036	0.003	0.021	0.08	0.59	21
Dassen	TripDuration	0.046	0.018	0.073	0.055	0.033	0.008	0.058	0.050	0.013	0.005	0.26	0.10	21
Robben	Condition	-0.009	-0.017	-0.001	0.015	-0.009	-0.017	-0.002	0.015	0.000	0.000	0.00	0.02	527
Robben	Growth	-0.002	-0.024	0.020	0.044	0.002	-0.019	0.024	0.043	-0.004	0.001	-0.09	0.03	56
Robben	MaxDist	0.004	-0.028	0.035	0.063	0.001	-0.020	0.021	0.041	0.003	0.022	0.07	0.54	27
Robben	PathLength	0.006	-0.024	0.035	0.058	-0.006	-0.024	0.012	0.036	0.011	0.022	0.31	0.61	25
Robben	TripDuration	0.006	-0.023	0.035	0.058	0.005	-0.021	0.031	0.052	0.001	0.006	0.02	0.11	26
Bird	Condition	-0.003	-0.018	0.012	0.030	-0.002	-0.017	0.012	0.029	-0.001	0.001	-0.03	0.02	207
Bird	MaxDist	0.010	-0.005	0.024	0.029	0.004	-0.009	0.018	0.027	0.005	0.002	0.20	0.08	57
Bird	PathLength	0.007	-0.010	0.025	0.035	0.012	-0.005	0.028	0.034	-0.004	0.001	-0.12	0.03	45
Bird	TripDuration	-0.019	-0.040	0.002	0.041	-0.014	-0.030	0.002	0.032	-0.005	0.010	-0.16	0.30	53
StCroix	Condition	0.002	-0.012	0.016	0.029	0.006	-0.008	0.02	0.028	-0.004	0.001	-0.143	0.024	135
StCroix	MaxDist	-0.036	-0.05	-0.022	0.028	-0.034	-0.048	-0.021	0.026	-0.001	0.002	-0.054	0.081	22
StCroix	PathLength	-0.026	-0.042	-0.009	0.034	-0.021	-0.037	-0.005	0.033	-0.005	0.001	-0.139	0.031	17
StCroix	TripDuration	0.000	-0.020	0.020	0.040	0.000	-0.015	0.016	0.031	0.000	0.009	-0.003	0.301	19

Table 4a: Estimates of change in population growth rates (expressed as an annual proportion) for the approaches in the Figure for the **West Coast**, and additionally model A1.5, which is the original aggregated model applied to the standardised data but with no outliers removed.

	Dassen						Robben					
	Approach	Estimate	se	Lower	Upper	Range	Model	Estimate	se	Lower	Upper	Range
Condition	A1	-0.002	0.004	-0.010	0.006	0.015	A1	-0.009	0.004	-0.017	-0.001	0.015
Condition	A1.5	-0.001	0.004	-0.009	0.007	0.016	A1.5	-0.008	0.004	-0.016	0.000	0.016
Condition	A2	0.001	0.004	-0.007	0.008	0.015	A2	-0.009	0.004	-0.017	-0.002	0.015
Condition	A3	0.001	0.004	-0.007	0.008	0.015	A3	-0.009	0.004	-0.017	-0.002	0.015
Condition	A1*	NA	NA	NA	NA	NA	A1*	NA	NA	NA	NA	NA
Condition	D0	-0.002	0.004	-0.010	0.006	0.016	D0	-0.009	0.004	-0.017	-0.002	0.016
Condition	D1	-0.002	0.004	-0.010	0.006	0.016	D1	-0.009	0.004	-0.017	-0.002	0.016
Condition	D2	0.001	0.004	-0.006	0.009	0.015	D2	-0.009	0.004	-0.017	-0.002	0.015
Condition	D3	0.000	0.004	-0.008	0.007	0.015	D3	-0.010	0.004	-0.018	-0.003	0.015
Growth	A1	0.018	0.010	-0.002	0.039	0.042	A1	-0.002	0.011	-0.024	0.020	0.044
Growth	A1.5	0.015	0.010	-0.006	0.035	0.041	A1.5	0.000	0.011	-0.021	0.022	0.043
Growth	A2	0.015	0.010	-0.006	0.035	0.041	A2	0.002	0.011	-0.019	0.024	0.043
Growth	A3	0.019	0.012	-0.006	0.043	0.049	A3	0.003	0.013	-0.023	0.029	0.052
Growth	A1*	0.024	0.013	-0.002	0.050	0.052	A1*	-0.002	0.014	-0.030	0.025	0.056
Growth	D0	0.021	0.014	-0.007	0.048	0.055	D0	0.001	0.014	-0.027	0.029	0.056
Growth	D1	0.021	0.014	-0.007	0.048	0.055	D1	0.001	0.014	-0.027	0.029	0.056
Growth	D2	0.020	0.015	-0.009	0.049	0.058	D2	0.000	0.015	-0.030	0.029	0.060
Growth	D3	0.020	0.013	-0.007	0.047	0.054	D3	0.001	0.013	-0.026	0.028	0.053
MaxDist	A1	0.004	0.015	-0.026	0.035	0.061	A1	0.004	0.016	-0.028	0.035	0.063
MaxDist	A1.5	0.004	0.016	-0.028	0.037	0.065	A1.5	0.004	0.017	-0.029	0.037	0.067
MaxDist	A2	-0.001	0.010	-0.021	0.019	0.040	A2	0.001	0.010	-0.020	0.021	0.041
MaxDist	A3	-0.008	0.011	-0.029	0.013	0.042	A3	-0.002	0.010	-0.022	0.018	0.040
MaxDist	A1*	NA	NA	NA	NA	NA	A1*	NA	NA	NA	NA	NA
MaxDist	D0	-0.006	0.007	-0.019	0.008	0.027	D0	-0.002	0.005	-0.013	0.008	0.021
MaxDist	D1	-0.002	0.011	-0.025	0.020	0.045	D1	0.002	0.011	-0.019	0.024	0.043
MaxDist	D2	-0.006	0.009	-0.024	0.012	0.036	D2	-0.004	0.008	-0.020	0.012	0.032
MaxDist	D3	-0.010	0.008	-0.026	0.005	0.031	D3	-0.002	0.007	-0.017	0.012	0.029
PathLength	A1	0.022	0.014	-0.007	0.050	0.057	A1	0.006	0.015	-0.024	0.035	0.058
PathLength	A1.5	0.029	0.015	-0.002	0.059	0.062	A1.5	0.003	0.016	-0.029	0.035	0.063
PathLength	A2	0.019	0.009	0.001	0.037	0.036	A2	-0.006	0.009	-0.024	0.012	0.036
PathLength	A3	0.013	0.010	-0.006	0.033	0.039	A3	-0.008	0.009	-0.026	0.010	0.036
PathLength	A1*	NA	NA	NA	NA	NA	A1*	NA	NA	NA	NA	NA
PathLength	D0	0.013	0.006	0.001	0.026	0.025	D0	-0.003	0.005	-0.013	0.007	0.020
PathLength	D1	0.016	0.011	-0.006	0.037	0.043	D1	-0.002	0.010	-0.023	0.018	0.041
PathLength	D2	0.007	0.008	-0.010	0.024	0.033	D2	-0.012	0.008	-0.027	0.003	0.030
PathLength	D3	0.005	0.008	-0.011	0.021	0.032	D3	-0.009	0.007	-0.023	0.004	0.027
TripDuration	A1	0.046	0.014	0.018	0.073	0.055	A1	0.006	0.014	-0.023	0.035	0.058
TripDuration	A1.5	0.048	0.016	0.015	0.081	0.066	A1.5	0.008	0.017	-0.026	0.042	0.068
TripDuration	A2	0.033	0.013	0.008	0.058	0.050	A2	0.005	0.013	-0.021	0.031	0.052
TripDuration	A3	0.031	0.013	0.004	0.057	0.053	A3	0.004	0.013	-0.022	0.030	0.052
TripDuration	A1*	NA	NA	NA	NA	NA	A1*	NA	NA	NA	NA	NA
TripDuration	D0	0.036	0.007	0.022	0.049	0.027	D0	0.003	0.005	-0.008	0.013	0.022
TripDuration	D1	0.038	0.012	0.013	0.062	0.049	D1	0.003	0.012	-0.020	0.027	0.047
TripDuration	D2	0.027	0.012	0.004	0.051	0.046	D2	-0.001	0.011	-0.023	0.020	0.042
TripDuration	D3	0.025	0.010	0.004	0.046	0.042	D3	0.001	0.010	-0.019	0.021	0.040

Table 4b: Estimates of change in population growth rates (expressed as an annual proportion) for the approaches in the Figure for the East Coast.

	Bird						St Croix					
	Approach	Estimate	se	Lower	Upper	Range	Model	Estimate	se	Lower	Upper	Range
Condition	A1	-0.003	0.007	-0.018	0.012	0.030	A1	0.002	0.007	-0.012	0.016	0.029
Condition	A1.5	-0.001	0.007	-0.015	0.012	0.027	A1.5	0.006	0.007	-0.008	0.019	0.026
Condition	A2	-0.002	0.007	-0.017	0.012	0.029	A2	0.006	0.007	-0.008	0.020	0.028
Condition	A3	-0.003	0.007	-0.017	0.012	0.029	A3	0.007	0.007	-0.007	0.020	0.027
Condition	A1*	NA	NA	NA	NA	NA	A1*	NA	NA	NA	NA	NA
Condition	D0	0.001	0.007	-0.014	0.016	0.030	D0	0.003	0.007	-0.011	0.018	0.029
Condition	D1	-0.004	0.007	-0.018	0.010	0.028	D1	0.003	0.007	-0.011	0.017	0.028
Condition	D2	-0.002	0.006	-0.015	0.010	0.025	D2	0.004	0.006	-0.008	0.017	0.025
Condition	D3	-0.004	0.007	-0.019	0.011	0.030	D3	0.003	0.007	-0.012	0.018	0.029
MaxDist	A1	0.010	0.007	-0.005	0.024	0.029	A1	-0.036	0.007	-0.050	-0.022	0.028
MaxDist	A1.5	0.005	0.007	-0.009	0.019	0.027	A1.5	-0.035	0.007	-0.048	-0.022	0.027
MaxDist	A2	0.004	0.007	-0.009	0.018	0.027	A2	-0.034	0.007	-0.048	-0.021	0.026
MaxDist	A3	0.002	0.006	-0.010	0.014	0.024	A3	-0.034	0.007	-0.047	-0.021	0.026
MaxDist	A1*	NA	NA	NA	NA	NA	A1*	NA	NA	NA	NA	NA
MaxDist	D0	0.004	0.003	-0.003	0.011	0.014	D0	-0.035	0.005	-0.046	-0.025	0.022
MaxDist	D1	0.008	0.006	-0.004	0.020	0.024	D1	-0.042	0.007	-0.056	-0.027	0.029
MaxDist	D2	0.004	0.006	-0.008	0.017	0.025	D2	-0.038	0.007	-0.052	-0.023	0.029
MaxDist	D3	0.007	0.006	-0.005	0.019	0.024	D3	-0.040	0.007	-0.054	-0.027	0.027
PathLength	A1	0.007	0.01	-0.01	0.025	0.035	A1	-0.026	0.01	-0.042	-0.009	0.034
PathLength	A1.5	0.01	0.01	-0.008	0.027	0.035	A1.5	-0.025	0.01	-0.042	-0.008	0.033
PathLength	A2	0.012	0.01	-0.005	0.028	0.034	A2	-0.021	0.01	-0.037	-0.005	0.033
PathLength	A3	0.009	0.01	-0.007	0.025	0.032	A3	-0.021	0.01	-0.037	-0.004	0.033
PathLength	A1*	NA	NA	NA	NA	NA	A1*	NA	NA	NA	NA	NA
PathLength	D0	-0.002	0	-0.01	0.006	0.016	D0	-0.025	0.01	-0.037	-0.012	0.025
PathLength	D1	0.005	0.01	-0.013	0.023	0.036	D1	-0.029	0.01	-0.049	-0.008	0.04
PathLength	D2	0.006	0.01	-0.011	0.024	0.035	D2	-0.024	0.01	-0.044	-0.004	0.04
PathLength	D3	0.005	0.01	-0.013	0.024	0.037	D3	-0.026	0.01	-0.044	-0.007	0.037
TripDuration	A1	-0.019	0.01	-0.04	0.002	0.041	A1	0	0.01	-0.02	0.02	0.04
TripDuration	A1.5	-0.017	0.01	-0.038	0.004	0.042	A1.5	0.001	0.01	-0.019	0.021	0.04
TripDuration	A2	-0.014	0.01	-0.03	0.002	0.032	A2	0	0.01	-0.015	0.016	0.031
TripDuration	A3	-0.014	0.01	-0.029	0	0.029	A3	0	0.01	-0.014	0.014	0.028
TripDuration	A1*	NA	NA	NA	NA	NA	A1*	NA	NA	NA	NA	NA
TripDuration	D0	-0.015	0	-0.022	-0.007	0.016	D0	-0.001	0.01	-0.013	0.012	0.025
TripDuration	D1	-0.007	0.01	-0.026	0.012	0.038	D1	-0.004	0.01	-0.024	0.017	0.042
TripDuration	D2	-0.006	0.01	-0.022	0.011	0.033	D2	-0.004	0.01	-0.023	0.015	0.037
TripDuration	D3	-0.006	0.01	-0.023	0.011	0.034	D3	-0.005	0.01	-0.021	0.011	0.033

Table 5: Sensitivity to period of data considered. Estimates and se's for the change in population growth rate (expressed as an annual proportion) for a variant of the A3 model, namely the standard island-closure GLM with sample-size-weighted and island-dependent variance, but applied to non-standardised data (note that for chick growth the lowest 12% of the data to exclude negative values were not removed here). This table includes results for fledging success, which is available in the form of annual aggregated values only. Results are shown for analyses of fledging success for data from 2008 onwards (corresponding to the years for which disaggregated foraging data are currently available), as well as for the entire data set. Additionally, results are shown for similar analyses of the chick growth, path length and trip duration response variables, for which pre-2008 data are available (these pre-2008 data are available in disaggregated form for chick growth, but only in aggregated form for the foraging data). The "All years' data" column essentially corresponds to analyses of the data presented in FISHERIES/2020/SEP/SWG-PEL/100. Note that where no disaggregated data are available to calculate the island-dependent variance in the same manner as for the other variables (see Appendix A), the standard deviations of the data points for each island were used instead. Note that for fledging success, when equivalent approaches were run in ADMB and R, the ADMB se's were a little larger (e.g. 0.18 in ADMB compared to 0.15 in R for the Robben island δ estimate for a variant of A3 with $\sigma_\epsilon=0$). This is most likely on account of insufficient information content in the data to partition variance reliably amongst the different components of the model. For other response variables the estimates from ADMB and R were very similar.

		Post 2008 only		All years' data		
		Estimate	se	Year range	Estimate	se
Condition	Dassen	-0.002	0.004	2004-2018	-0.002	0.004
	Robben	-0.009	0.004	2004-2018	-0.006	0.004
	Bird	-0.004	0.008	2008-2018	Available only from 2008	
	StCroix	0.003	0.007	2008-2018		
Growth	Dassen	0.024	0.013	1989-2014	0.036	0.012
	Robben	-0.002	0.014	2004-2014	0.003	0.013
MaxDist	Dassen	0.001	0.015	2008-2018	Available only from 2008	
	Robben	0.002	0.015	2008-2018		
	Bird	0.009	0.007	2008-2018		
	StCroix	-0.038	0.007	2008-2018		
PathLength	Dassen	0.020	0.013	2003-2018	0.033	0.015
	Robben	0.003	0.013	2003-2018	0.002	0.015
	Bird	0.006	0.008	2007-2018	0.003	0.010
	StCroix	-0.028	0.008	2008-2018	-0.018	0.010
TripDuration	Dassen	0.044	0.014	2003-2018	0.047	0.013
	Robben	0.005	0.014	2003-2018	0.010	0.013
	Bird	-0.020	0.009	2007-2018	-0.016	0.011
	StCroix	-0.001	0.009	2008-2018	-0.003	0.011
Fledging success	Dassen	-0.008	0.008	1995-2015	0.008	0.014
	Robben	0.004	0.010	1989-2015	-0.015	0.016

Table 6: Sensitivity to exclusions of multiple trips by the same bird. Data rows with TripID=1 only were retained (i.e. rows with a TripID=2,3,...etc which correspond to the 2nd, 3rd etc trip for a particular bird were excluded) before aggregating and applying the A3 model. Estimates and se's for the change in population growth rate (expressed as an annual proportion) are listed for: (a) the original post—2008 data with no removals of multiple trips; and (b) the re-run with the removals. The third column lists the difference between the two.

		(a) All post 2008 data		(b) Repeat BirdIDs removed		(b)-(a)	
		Estimate	se	Estimate	se	Estimate	se
MaxDist	Dassen	0.001	0.015	0.004	0.015	0.003	0.000
	Robben	0.002	0.015	0.001	0.014	-0.001	-0.001
	Bird	0.009	0.007	0.008	0.006	-0.001	-0.001
	StCroix	-0.038	0.007	-0.036	0.007	0.002	0.000
PathLength	Dassen	0.020	0.013	0.025	0.014	0.005	0.001
	Robben	0.003	0.013	0.003	0.013	0.000	0.000
	Bird	0.006	0.008	0.006	0.009	0.000	0.001
	StCroix	-0.028	0.008	-0.025	0.009	0.003	0.001
TripDuration	Dassen	0.044	0.014	0.049	0.013	0.005	-0.001
	Robben	0.005	0.014	0.003	0.012	-0.002	-0.002
	Bird	-0.020	0.009	-0.009	0.010	0.011	0.001
	StCroix	-0.001	0.009	0.003	0.010	0.004	0.001

Table 7a: Sensitivity to estimation method. Estimates and se's for the change in population growth rate (expressed as an annual proportion) for approaches A2 and D3 for estimation based on MLE (i.e. as for the previous results), as well as on REML. Note that model A2, rather than A3, was used, as model A3 is coded in ADMB, for which there is not a REML option.

			A2		D3	
			Estimate	se	Estimate	se
Condition	Dassen	MLE	0.001	0.004	-0.001	0.004
		REML	0.001	0.004	-0.002	0.004
	Robben	MLE	-0.009	0.004	-0.010	0.004
		REML	-0.009	0.004	-0.009	0.004
	Bird	MLE	-0.002	0.007	-0.004	0.007
		REML	-0.002	0.008	-0.004	0.008
	StCroix	MLE	0.006	0.007	0.003	0.007
		REML	0.006	0.008	0.003	0.008
Growth	Dassen	MLE	0.015	0.010	0.020	0.013
		REML	0.015	0.012	0.021	0.017
	Robben	MLE	0.002	0.011	0.001	0.013
		REML	0.002	0.013	0.001	0.017
MaxDist	Dassen	MLE	-0.001	0.010	-0.010	0.008
		REML	-0.001	0.011	-0.001	0.013
	Robben	MLE	0.001	0.010	-0.002	0.007
		REML	0.001	0.011	0.003	0.012
	Bird	MLE	0.004	0.007	0.007	0.006
		REML	0.004	0.008	0.008	0.007
	StCroix	MLE	-0.034	0.007	-0.040	0.007
		REML	-0.035	0.007	-0.042	0.008
PathLength	Dassen	MLE	0.019	0.009	0.005	0.008
		REML	0.019	0.010	0.016	0.012
	Robben	MLE	-0.006	0.009	-0.009	0.007
		REML	-0.006	0.010	-0.002	0.012
	Bird	MLE	0.012	0.008	0.005	0.009
		REML	0.012	0.009	0.005	0.010
	StCroix	MLE	-0.021	0.008	-0.026	0.009
		REML	-0.021	0.009	-0.028	0.011
TripDuration	Dassen	MLE	0.033	0.013	0.025	0.010
		REML	0.033	0.014	0.038	0.014
	Robben	MLE	0.005	0.013	0.001	0.010
		REML	0.005	0.015	0.003	0.013
	Bird	MLE	-0.014	0.008	-0.006	0.009
		REML	-0.014	0.009	-0.007	0.011
	StCroix	MLE	0.000	0.008	-0.005	0.008
		REML	0.000	0.009	-0.003	0.012

Table 7b: Summary of the differences in the estimates and se's for changes in population growth rates (expressed as an annual proportion) for the REML and MLE estimation methods in

Table 7a. The bottom section summarises the differences averaged as appropriate by response variable type (condition, growth, as well as all the foraging data combined) and coast.

		se(REML)-se(MLE)		se(REML)/se(MLE)	
		A2	D3	A2	D3
Condition	Dassen	0.0004	0.0005	1.106	1.144
	Robben	0.0004	0.0005	1.106	1.141
	Bird	0.0007	0.0004	1.104	1.059
	StCroix	0.0007	0.0004	1.105	1.059
Growth	Dassen	0.0020	0.0031	1.201	1.231
	Robben	0.0022	0.0039	1.202	1.295
MaxDist	Dassen	0.0011	0.0051	1.110	1.653
	Robben	0.0011	0.0052	1.111	1.712
	Bird	0.0007	0.0007	1.104	1.109
	StCroix	0.0007	0.0010	1.104	1.146
PathLength	Dassen	0.0010	0.0043	1.110	1.538
	Robben	0.0010	0.0052	1.111	1.770
	Bird	0.0009	0.0009	1.106	1.095
	StCroix	0.0009	0.0018	1.106	1.192
TripDuration	Dassen	0.0014	0.0031	1.111	1.302
	Robben	0.0015	0.0032	1.111	1.326
	Bird	0.0009	0.0022	1.107	1.255
	StCroix	0.0008	0.0033	1.106	1.407
Averages					
Condition	WC	0.0004	0.0005	1.106	1.142
	EC	0.0007	0.0004	1.104	1.059
Growth	WC	0.0021	0.0035	1.201	1.263
Foraging	WC	0.0012	0.0044	1.111	1.550
	EC	0.0007	0.0014	1.105	1.029

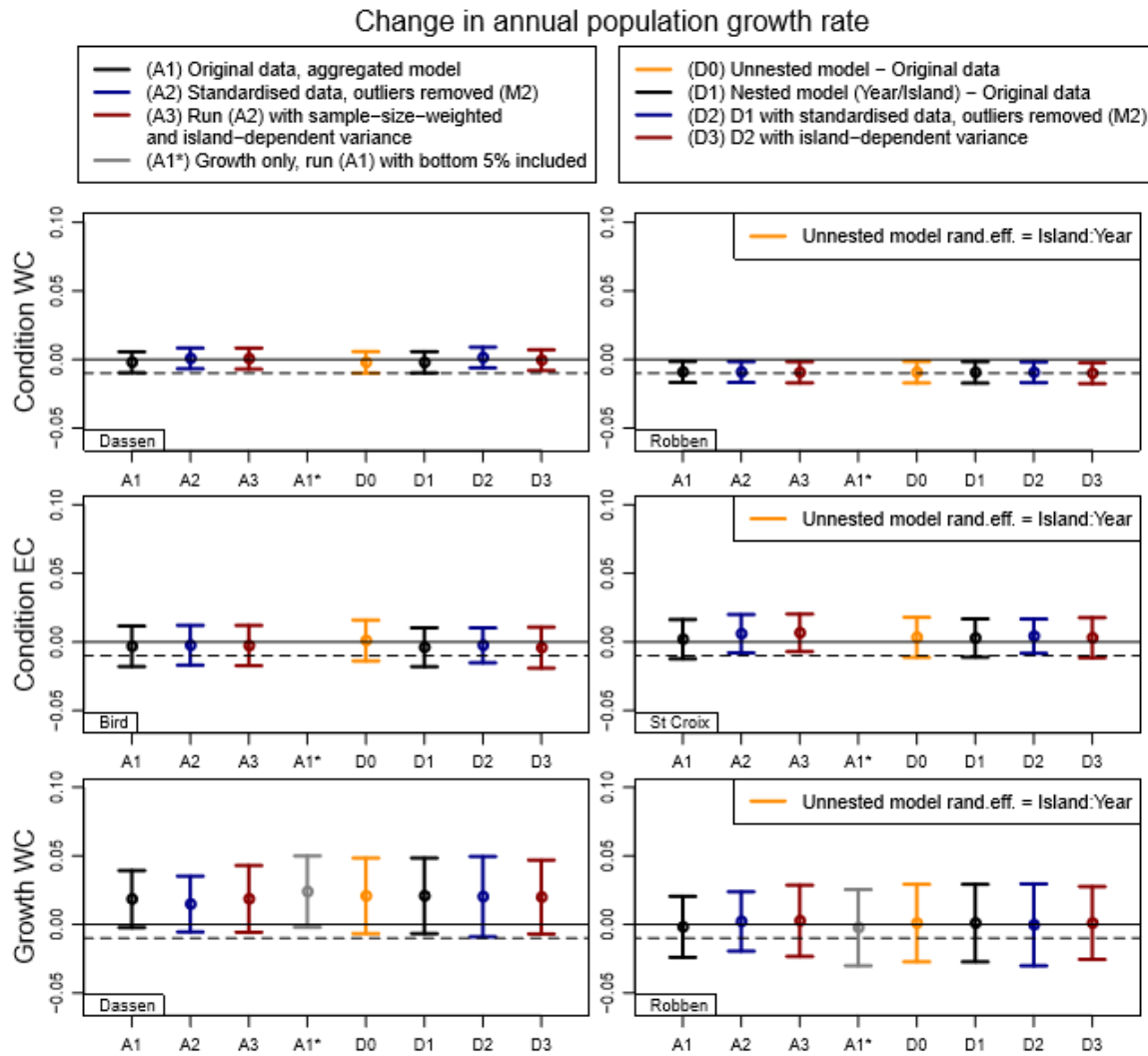


Figure 1(a): Estimates and 95% confidence intervals (given by \pm twice the standard error of the estimates) of the change in annual population growth rate (expressed as an annual proportion) for **chick condition** (WC and EC) and **chick growth rate** (WC only) are shown for a range of aggregated and disaggregated approaches, applied to original and standardised data. These estimates are derived from the island-closure GLM δ estimates by multiplying the δ estimates and se's by 0.094 (for all variables other than chick condition) and 0.11 for chick condition (REF).

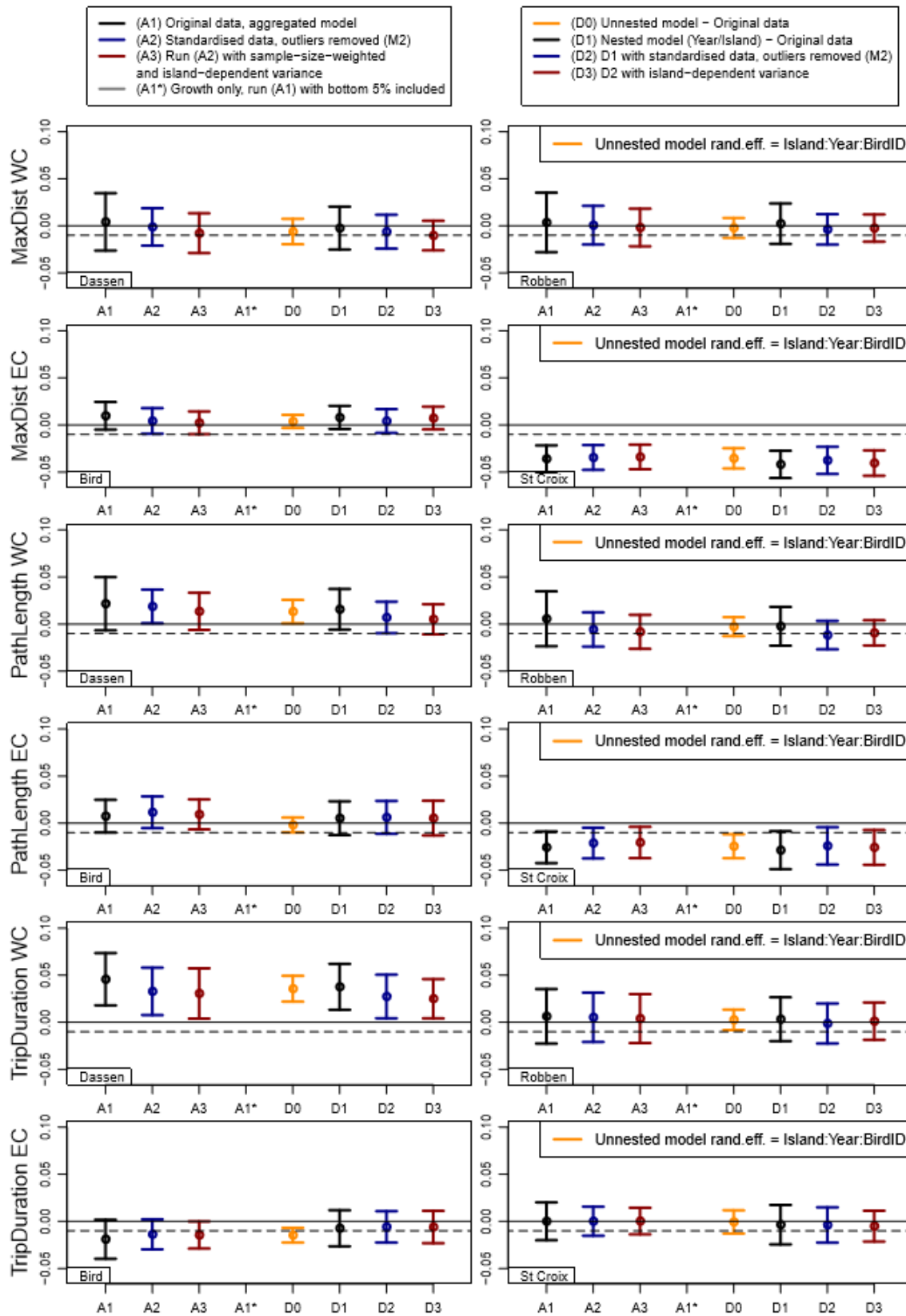


Figure 1(b): Estimates and 95% confidence intervals of the change in annual population growth rate (expressed as an annual proportion) for the **foraging data**, on a similar basis to Figure 1(a).

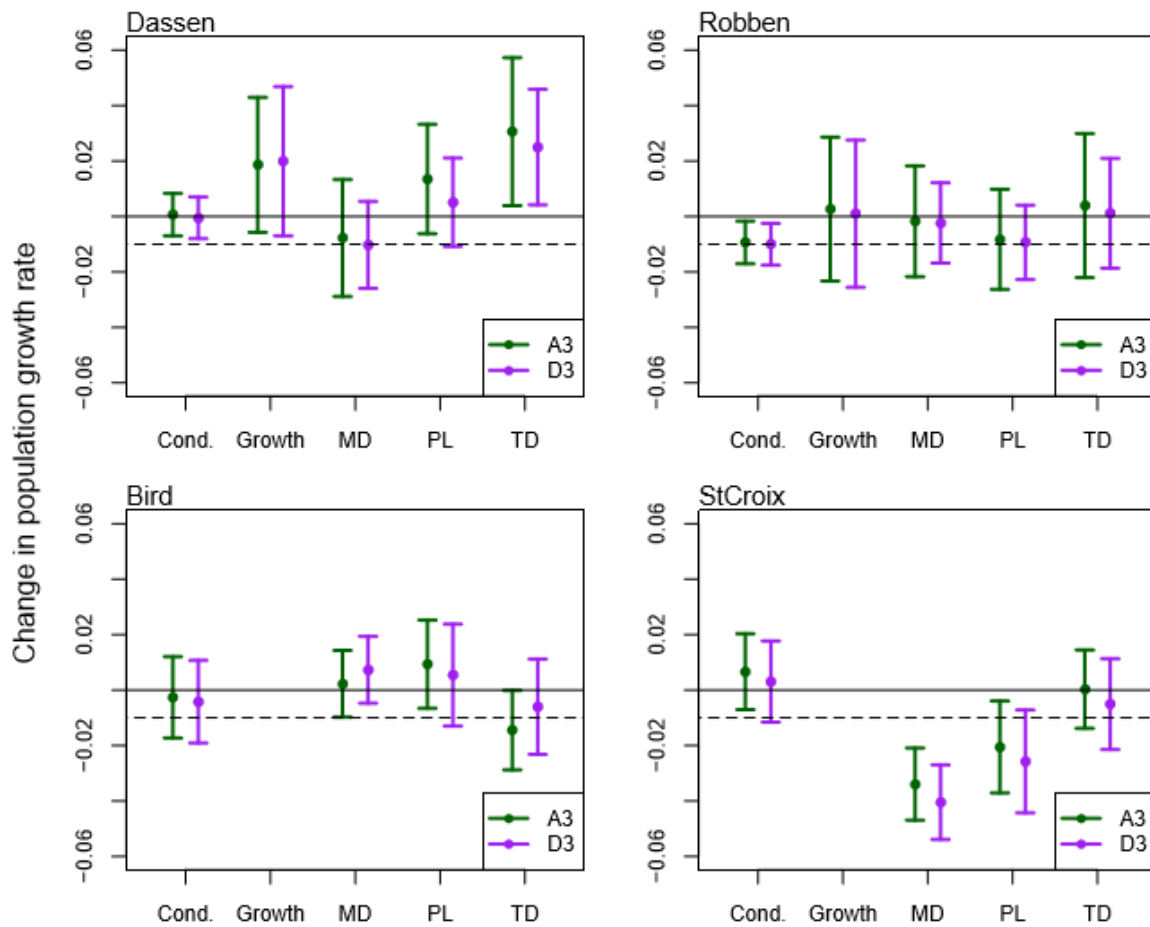


Figure 2: Comparison of the estimated change in annual population growth rate (expressed as an annual proportion) for approaches **A3** and **D3** to compare the “best” implementations of the aggregated and disaggregated approaches. Point estimates and 95% CIs are shown for the various response variables for which month information is available. In the axis labels, “MD” is maximum foraging distance, “PL” is path length and “TD” is trip duration.

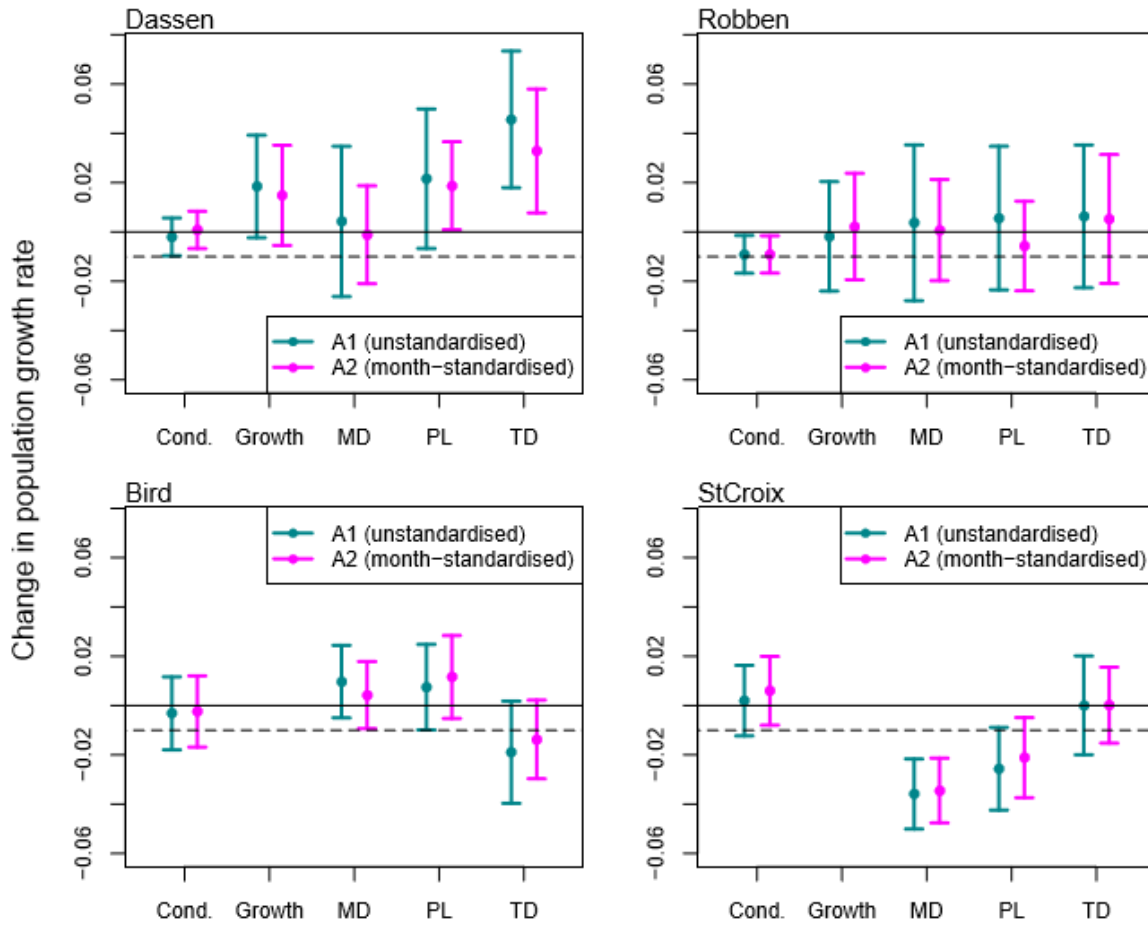


Figure 3: Comparison of the estimated change in annual population growth rate (expressed as an annual proportion) for approaches **A1** and **A2** to show the impact of month-standardisation. Point estimates and 95% CIs are shown for the various response variables for which month information is available. In the axis labels, “MD” is maximum foraging distance, “PL” is path length and “TD” is trip duration.

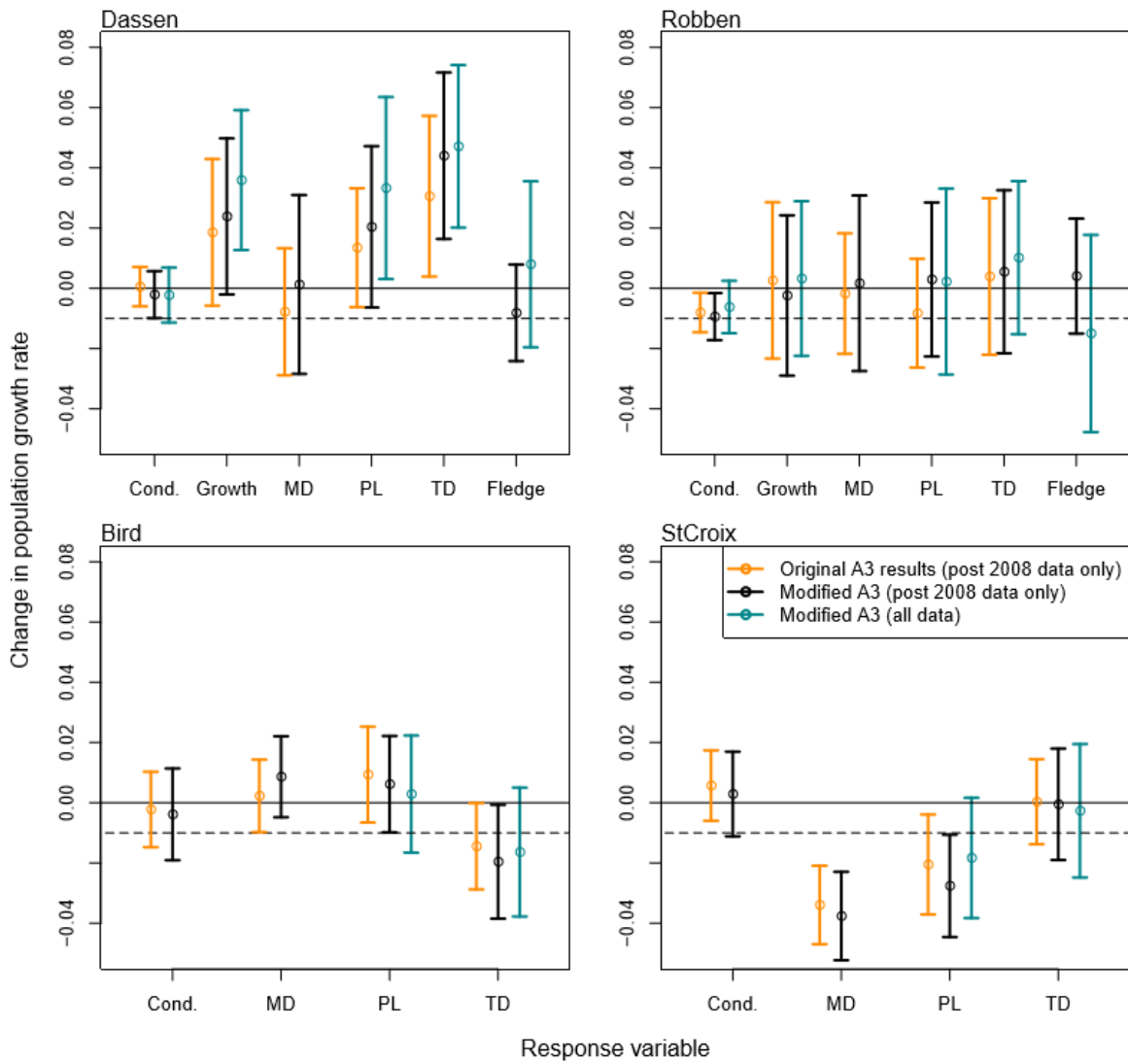


Figure 4: Sensitivity to data period considered. Zeh plots of the estimates of change in population growth rate (expressed as an annual proportion) from **Table 5**, for a variant of the A3 model, namely the standard island-closure GLM with sample-size-weighted and island-dependent variance applied to non-standardised aggregated data. Results are shown for all response variables except chick survival. The black points and error bars show the results when data from 2008 onwards only are analysed, while the turquoise points and error bars show these when data prior to 2008 (where these are available) are included. The original A3 results have been included in orange (i.e. the same results as the dark green points in Figure 2). The difference between the orange and the black bars is that the orange use standardised data, while the black do not.

Appendix A: Extracts from FISHERIES/2021/APR/SWG-PEL/24²

An initial implementation of suggestions by the 2020 Panel to improve estimates of the effect of fishing around islands on penguins by using models with random effects and applied to both aggregated and disaggregated data

Summary

Pending availability of data on the month for each observation for further analyses of the island closure experiment to address the suggestions made by the 2020 Panel, some initial analyses are conducted to explore the implications of the Panel's suggestions regarding sample size weighting for data-aggregated approaches, and nesting of random effects for methods using disaggregated data. These are applied to existing data for maximum foraging distance and chick condition. Results indicate that when these Panel suggestions are taken into account, there is little to no real difference in the results which these two approaches provide for the effect of fishing parameter. This is as the Panel anticipated, and in line with an earlier mathematical demonstration that use of the extra data available under the disaggregated approach would not improve the precision of these estimates. Some corresponding earlier results suggesting greater precision for the disaggregated approach were a consequence of applying a random effects approach without appropriate nesting of the data – an approach which is rejected by statistical analyses. Results for estimates of the effect of fishing parameter are more meaningfully expressed in terms of the consequent impact on the penguin population growth rate. For the cases examined initially here, the only meaningful result suggesting a negative impact of fishing in the vicinity of islands is for the maximum foraging distance variable for St Croix island.

Introduction

The 2020 Panel (Haddon *et al.* 2020) made a number of suggestions to advance the estimation of the effects of fishing around islands with colonies of penguins, particularly as regards the comparison of approaches applied to either annually aggregated or dis-aggregated (individual) data. These included:

- Use of common data and a common model structure for comparisons
- Inclusion of month as a co-variate
- Consideration of different nesting structures for models using random effects.

Finalised data with information on the month for each observation have been in the process of being compiled.

Methods

Question 1: How does the precision of the estimated fishing effect differ when models including random effects are applied to aggregated compared to disaggregated data?

A range of random effects models have been applied to the maximum foraging distance and chick condition data sets, with these data both in aggregated form (aggregated to yield a single value per year and island) and disaggregated form. The basic equation for these mixed models is:

² For reference purposes the same table and figure numbering has been retained as in the original document. As not all tables and figures have been retained from the original document, there are therefore some gaps in the numbering.

$$f(F_{y,i}) = R + \beta_i + \delta_i^* X_{i,y} + \epsilon_{y,i,s} \quad (1)$$

where

$F_{y,i}$	is the response variable, possibly log transformed (see below for more details for year y and island i),
R	represents the random effects components (more details below),
β_i	is an island effect,
δ_i^*	is a fishing effect resulting from whether or not the neighbourhood around the island was closed to fishing ³ ,
X_i	is a vector with an entry of 1 in years where the neighbourhood of island i was open to the fishery, and 0 where that neighbourhood was closed, and
$\epsilon_{y,i}$	is an error term, taken to be normally distributed with constant variance unless otherwise indicated.

Data transformation

As has been standard practice, the maximum foraging distance data have been log transformed by taking the negative of the logarithm so that a larger negative number implies a more negative impact on the penguin population, i.e. $f(F_{y,i}) = -\ln F_{y,i}$. For chick condition data, however, no transformation has been applied, *inter alia* because of negative values in the disaggregated data set given the manner in which this variable is defined, so that $f(F_{y,i}) = F_{y,i}$, and again a negative number implies a negative impact on the penguin population

Random effects components

When applied to aggregated data (both maximum distance and chick condition), the random effects component of equation (1) is simply (1|Year), i.e. a random year effect reflecting prevailing feeding conditions (assumed to be the same each year, random variation excepted, for both islands in a pair – because of their close proximity, these conditions at island should each relate closely to the forage fish abundance for that year).

When applied to the disaggregated data, a key question is whether nested or unnested models should be used. It is our understanding that Sherley et al. (2018), albeit unintentionally, used unnested models, while the general preference and particularly that of the review Panel is for nested models (Sherley 2020a&b, Haddon et al. 2020, Bergh 2021).

To illustrate the difference between a nested and an unnested model, consider a mixed model with a Year-Island random effects component. In R notation, an unnested model would have the following form:

`lmer(lnResponse ~ (1|Year:Island) + Island + X:Island)` (Model I)

However, a nested model, with island nested within year, has the notation:

`lmer(lnResponse ~ (1|Year/Island) + Island + X:Island)` (Model IIa)

This Model IIa is exactly the same as:

`lmer(lnResponse ~ (1|Year) + (1|Year:Island) + Island + X:Island)` (Model IIb)

Nesting island within year implies that response variables for the two islands are correlated from year to year (i.e. a common (1|Year) component in addition to the island specific (1|Year:Island) component), whereas an unnested model

³ Note that this equation has been modified for greater generality depending to allow for different transformations (f) of the response variable (F). Accordingly, the fishing effect parameter has been generalised to δ^* . For maximum foraging distance, where the data have been log-transformed, δ^* is the standard δ parameter for which results have been reported in previous analyses. For chick condition, where the data have not been transformed, the estimated value of what is now termed the $\tilde{\delta}$ parameter needs to be interpreted differently from what has been done in the past (e.g. with respect to the assumed threshold below which fishing has a biologically meaningful impact on the penguin population) – see the Appendix B for more details.

implies that they are not (there is only the island specific (1|Year:Island) random effects component). The rationale underlying the design of the experiment with different treatments of the two islands from year to year was based on the assumption that such correlation exists (because the proximity of the pair of islands would mean that each year the forage fish densities near each island would tend to vary in the same direction about their averages over time), and provides a basis to separate the otherwise confounded effects of year and closure, thereby providing better estimation precision.

Note that the unnested Model I is essentially a “sub-model” of the nested Model II, so that a likelihood ratio test can be used to compare the fits of the two. Results for such tests are provided in this paper to provide an objective quantitative basis for deciding whether a nested or unnested model is to be preferred.

A second element is the inclusion in the random effects components of lower level information such as BirdID in the case of maximum foraging distance, and month for chick condition. Models including and excluding these further levels of information have also been explored.

Taking sample size into account for the aggregated model

One criticism of the aggregated approach as applied in the past (Haddon et al. 2020) is that it does not take into account the number of data points used to calculate the annual averages. In order to investigate this (an option not present in R), the equation (1) mixed model (applied to the aggregated data) was recoded in AD Model Builder (ADMB) with the variance for the model residuals redefined as:

$$\sigma_{\epsilon}^2(i, y) = \frac{\sigma_{i,y}^2}{N_{i,y}} + \tau^2 \quad (1)$$

where $\sigma_{i,y}$ is the standard deviation of the data collected for island i in year y , $N_{i,y}$ is the number of samples collected for island i in year y , and τ^2 is the remaining residual variance and an estimable parameter. On inspection of the annual sd values, it was decided that a sample-size-weighted average over years, σ^ , would be a more robust approach because the precision of year-specific estimates is poor because of instances of low sample size, i.e.:*

$$\text{island-combined: } (\sigma^*)^2 = \sum_{i,y} \sigma_{i,y}^2 (N_{i,y} - 1) / \sum_{i,y} (N_{i,y} - 1)$$

$$\text{island-specific: } (\sigma_i^*)^2 = \sum_y \sigma_{i,y}^2 (N_{i,y} - 1) / \sum_y (N_{i,y} - 1)$$

Results for σ^ are listed in Table 2 of FISHERIES/2021/APR/SWG-PEL/24.*

Question 2: What is the threshold value for the island closure effect parameter which corresponds a (“biologically meaningful”) 1% change in the annual population growth rate?

In analyses of penguin data to date, estimates of the fishing effect have been taken to be biologically meaningful if they correspond to a change in the population growth rate of more than 1% pa. If response variables analysed are considered in log space, this has been assumed to correspond to a value of the fishing effect parameter δ which is < -0.1 (see Robinson *et al.* 2014 and the Appendix below) using the relationship between annual survival and the population growth rate. The question is whether this assumption can be replaced by a relationship established from data for at least some of the other response variables. To attempt to answer this question for chick condition, the data available are analysed here using the following approach: (1) establish a relationship between chick condition and annual survival for another penguin population, and (2) substitute this relationship into the demographic equations of Robinson *et al.* (2014) (which correspond to the corresponding Leslie matrix eigenvalue equation) to establish an appropriate threshold for a change in the value of chick condition corresponding to a 1% change in the annual penguin population growth rate. Further details for this analysis are provided in the Appendix B.

Results

Table 3 gives the details of all the models that have been implemented in an attempt to answer Question 1. Table 4(a) lists the results for these models applied to the West Coast maximum distance foraging trip data. For all results presented in this Appendix, the models were run with the Maximum Likelihood Estimation (MLE) method. While REML would be the preferred approach for all models, difficulties were encountered in implementing REML for the ADMB models, and therefore in the interest of comparability all the main models in Table 4(a) were also run with MLE. Table

5a to Table 7a list the corresponding results for EC maximum distance, WC chick condition and EC chick condition respectively.

In most cases, the precision for δ^* from the aggregated and disaggregated approach were fairly similar⁴. The one notable exception was when an unnested model that included BirdID as a covariate in the random effects component was applied to the maximum distance foraging data. This model yielded a standard error estimate for δ^* which was roughly half that estimated by nested models and models applied to aggregated data.

Table 8 lists the results for the likelihood ratio test conducted on a selection of models to investigate whether nested or unnested models are to be preferred. In many cases the nested model is not significantly better than the unnested model, but importantly for maximum distance when BirdID is included in the random effects structure there is strong statistical support for the nested model.

The analyses of Appendix B indicate that a $\tilde{\delta}$ threshold of -0.09 for chick condition (when these data are not log-transformed) corresponds to a 1% reduction in population growth rate.

Discussion

Question 1: How does the precision of the estimated fishing effect δ differ when models including random effects are applied to aggregated compared to disaggregated data?

- In most cases, the precisions (se' s) estimated for δ^* using the aggregated and disaggregated approaches are in fact fairly similar. The only model with a substantially lower se is D0_unnested applied to the maximum distance foraging data, i.e. an unnested model including BirdID. The likelihood ratio tests (Table 8) show that when BirdID is included, there is clear statistical preference for the nested model, and therefore there is an objective basis to disregard the unnested models and their smaller estimates of the δ standard error (in addition to views expressed to prefer the nested models for other reasons).
- Inclusion of lower level information such as BirdID and month does not make a substantial difference provided the model has an appropriate hierarchical (nested) structure.

Question 2: How suitable is the value of -0.1 as the threshold for δ for a biologically meaningful impact on the penguin population?

- For the chick condition response variable, negative values preclude simple use of the customary log-transformation approach because of the presence of negative individual values. However, there is a relationship available for another penguin population (Macaroni penguins, see Sherley *et al.* 2018) which allows a change in chick condition to be related to a change in penguin population growth rate.
- For a “biologically meaningful” change of 1% in penguin population growth rate requires a change of δ of 0.106 (the critical threshold value, inverse of Equation A9 of Appendix B) for the standard log transformation of the response variable (as for maximum foraging distance), but for the specific case of chick condition with untransformed data the corresponding value of $\tilde{\delta}$ is 0.091 (inverse of Equation A10 of Appendix B).
- For comparison of results for different response variables, it is preferable to show these in terms of the estimated change in penguin population growth rate.

Other points worth noting include:

- Where models have been recoded in ADMB, estimates of precision for the random effects' standard deviation and residuals standard deviations can be obtained. These are shown in parenthesis in Tables 4-7. The variance of the sd parameters are fairly large, indicating that the partitioning of variance to different sources (e.g. random effects vs model residuals) is not able to be achieved very well for these data sets.

⁴ δ^* refers to the conventional δ for the log-transformed maximum distance data, and to $\tilde{\delta}$ for the untransformed chick condition data.

References

- Bergh, M. 2021. An initial response to FISHERIES/2021/FEB/SWG-PEL/08. DEFF Fisheries document FISHERIES/2021/MAR/SWG-PEL/013.
- Haddon, M., Parma, A., Punt, A.E. and Wilberg, M.J. 2020. Report of the international review of some aspects of the island closure experiment. DEFF Fisheries document FISHERIES/2020/DEC/SWG-PEL/REVIEW/07.
- Robinson, W.M.L., Butterworth, D.S. and Furman, L.B. 2014. Analyses of the Results from the Island Closure Feasibility Study for the Dassen/Robben and St Croix/Bird Island Pairs. International Stock Assessment Workshop document MARAM/IWS/DEC14/Peng/B4.
- Sherley, R.B. *et al.* 2018. Bayesian inference reveals positive but subtle effects of experimental fishery closures on marine predator demographics. *Proceedings of the Royal Society B: Biological Sciences*. <https://doi.org/10.1098/rspb.2017.2443>
- Sherley, R.B. 2020a. Revisiting the key results in MARAM/IWS/2019/PENG/P4 in light of the 2019 Panel recommendations. DEFF Fisheries document FISHERIES/2020/JUL/SWG-PEL/53.
- Sherley, R.B. 2020b. Some observations on comparisons of fitting to the annually aggregated and the individual data, this time using JAGS and for the cases considered in FISHERIES/2020/JUL/SWG-PEL/53REV. DEFF fisheries document FISHERIES/2020/SEP/SWG-PEL/86.

Table 3: Descriptions of the models for which results have been presented in Tables 4-7, and the R code used to implement them. Note that the “BirdID” co-variate is relevant for the maximum foraging distance data sets; for the chick condition analyses “BirdID” is replaced by “month”. The “lmer” function is from the R lme4 package, and the “lme” function from the R nlme package.

Aggregated data	Description	R Code
A1	The original mixed model (estimates are MLE)	<code>lmer(lnResponse ~ (1 Year) + Island + X:Island, data=Data, REML=FALSE)</code>
A2	A1, but with unequal variance between the islands.	<code>lme(lnResponse ~ Island + X:Island, random=~1 Year, weights=varIdent(form=~1 Island), data=Data, method="ML")</code>
A1+N	A1, with sample-size adjusted variance	Coded in AD Model Builder (ADMB), with $\sigma_{\epsilon}^2(i, y) = \frac{(\sigma_{iy}^*)^2}{N_{i,y}} + \tau^2$
A2+N	A2, with sample-size adjusted variance	Coded in AD Model Builder (ADMB), with $\sigma_{\epsilon}^2(i, y) = \frac{(\sigma_{iy}^*)^2}{N_{i,y}} + \tau^2$
Disaggregated data	Description	R Code
D0_unnest	the M1 models of Table 2 of FISHERIES/2020/JUL/SWG-PEL/53 (NOT nested)	<code>lmer(lnResponse ~ (1 Island:Year:BirdID) + Island + X:Island, data=Data, REML=FALSE)</code>
D1_unnest	the M4 models of Table 2 of FISHERIES/2020/JUL/SWG-PEL/53 (NOT nested)	<code>lmer(lnResponse ~ (1 Island:Year) + Island + X:Island, data=Data, REML=FALSE)</code>
D0	D0 with nesting included	<code>lmer(lnResponse ~ (1 Year/Island/BirdID) + Island + X:Island, data=Data, REML=FALSE)</code>
D1	D1 with nesting included	<code>lmer(lnResponse ~ (1 Year/Island) + Island + X:Island, data=Data, REML=FALSE)</code>
D0+uneqvar	D0 with unequal variance for the islands	<code>lme(lnResponse ~ Island + X:Island, random=~1 Year/Island/BirdID, weights=varIdent(form=~1 Island), data=Data, method="ML")</code>
D1+uneqvar	D1 with unequal variance for the islands	<code>lme(lnResponse ~ Island + X:Island, random=~1 Year/Island, weights=varIdent(form=~1 Island), data=Data, method="ML")</code>

Table 4a: Various models (see Table 3 for descriptions of the models) are applied to the aggregated and disaggregated **West Coast maximum distance foraging** data. Estimates and se's for the closure effect δ are listed (including the average se across the two islands), as well as the random effects and residuals standard deviations. Where several values are listed under the random effects' header, these correspond to the different random components of the model. e.g. Year/Island/Bird (D0) shows the sd for Bird:(Island:Year) then Island:Year then Year. The residuals standard deviation has been obtained in two ways. RES_sd1 is the standard R output as provided by the VarCorr function for the lmer (random effects) application. RES_sd2 is the standard deviation of the residuals, calculated as $sd(y-\hat{y})$ where y is $-\ln(\text{data})$ and $\hat{y}=\text{predict}(\text{Fit})$. For the models with unequal variance, the residual standard deviation RES_sd1 is given first for Dassen, then for Robben. For A1+N and A2+N (where residual variance has been defined as $\sigma_{\epsilon}^2(i,y) = \frac{\sigma_{\epsilon,y}^2}{N_{i,y}} + \tau^2$, the RES_sd1 column reports the τ estimates. Where models have been recoded in ADMB, estimates of precision for the random effects' standard deviation and residuals standard deviations can be obtained. Where such information is available, it has been shown in parenthesis.

	Description		Dassen		Robben		ave(se)	Random effects sd			Residuals sd	
		Random effects	delta	se	Delta	se		Year,Island,Bird	Year,Island	Year (σ_{α})	RES_sd1 (σ_{RES})	RES_sd2
Aggregated data												
A1	Original model	Year	0.045	0.162	0.039	0.168	0.165	-	-	0.160 (0.076)	0.221 (0.049)	0.195
A2	Unequal variance for the islands	Year	0.028	0.206	0.038	0.151	0.179	-	-	0.161	0.269 & 0.152	0.200
A1+N	A1, with sample-size adjusted variance	Year	0.008	0.158	0.018	0.152	0.155	-	-	0.152 (0.065)	0.170 (0.054)	0.199
A2+N	A2, with sample-size adjusted variance	Year	0.013	0.158	0.018	0.155	0.157			0.152 (0.067)	0.175 (0.054)	0.199
Disaggregated data												
D0_unnest	Island:Year:Bird	Island:Year:Bird	-0.064	0.072	-0.020	0.056	0.064	0.343	-	-	0.291	0.196
D1_unnest	Island:Year	Island:Year	-0.005	0.127	0.025	0.120	0.124	-	0.167	-	0.426	0.420
D0	Year/Island/Bird	Year/Island/Bird	-0.025	0.116	0.010	0.108	0.112	0.313	0.095	0.119	0.291	0.202
D1	Year/Island	Year/Island	-0.025	0.121	0.025	0.114	0.118	-	0.120	0.110	0.427	0.421
D0+uneqvar	D0 with unequal variance for the islands	Year/Island/Bird	-0.023	0.116	0.011	0.109	0.113	0.289	0.097	0.119	0.289 & 0.299	0.208
D1+uneqvar	D1 with unequal variance for the islands	Year/Island	-0.019	0.122	0.025	0.118	0.120	-	0.111	0.124	0.403 & 0.445	0.420

Table 5a: Repeat of Table 4a for **East Coast maximum distance foraging** data.

	Description	Random effects	St Croix		Bird		ave(se)	Random effects sd			Residuals sd	
			delta	se	delta	se		Year,Island,Bird	Year,Island	Year	RES_sd1	RES_sd2
Aggregated data												
A1	Original model	Year	-0.367	0.072	0.088	0.075	0.074	-	-	0.072 (0.033)	0.098 (0.021)	0.086
A2	Unequal variance for the islands	Year	-0.366	0.088	0.086	0.072	0.080	-	-	0.072	0.114 & 0.078	0.087
A1+N	A1, with sample-size adjusted variance	Year	-0.393	0.079	0.073	0.066	0.072	-	-	0.083 (0.027)	0.031 (0.047)	0.084
A2+N	A2, with sample-size adjusted variance	Year	-0.388	0.075	0.075	0.069	0.072	-	-	0.081 (0.027)	0.044 (0.036)	0.082
Disaggregated data												
D0_unnest	Island:Year:Bird	Island:Year:Bird	-0.377	0.057	0.040	0.037	0.047	0.209	-	-	0.379	0.337
D1_unnest	Island:Year	Island:Year	-0.419	0.085	0.076	0.073	0.079	-	0.101		0.423	0.419
D0	Year/Island/Bird	Year/Island/Bird	-0.440	0.075	0.081	0.063	0.069	0.188	0.000	0.091	0.378	0.340
D1	Year/Island	Year/Island	-0.445	0.077	0.085	0.065	0.071	-	0.000	0.097	0.423	0.421
D0+uneqvar	D0 with unequal variance for the islands	Year/Island/Bird	-0.440	0.070	0.082	0.064	0.067	0.201	0.000	0.092	0.305 & 0.394	0.330
D1+uneqvar	D1 with unequal variance for the islands	Year/Island	-0.446	0.072	0.085	0.065	0.069	-	0.000	0.097	0.364 & 0.443	0.421

Table 6a: Various models are applied to the aggregated and disaggregated **West Coast chick condition** data. Quantities are as defined in the caption to Table 4a, except that here the data have not been log transformed.

	Description		Dassen		Robben		ave(se)	Random effects sd			Residuals sd	
		Random effects	delta	se	Delta	se		Year,Island,Month	Year,Island	Year (σ_{α})	RES_sd1 (σ_{RES})	RES_sd2
Aggregated data												
A1	Original model	Year	-0.018	0.035	-0.083	0.035	0.035	-	-	0.016 (0.031)	0.055 (0.012)	0.055
A2	Unequal variance for the islands	Year	-0.018	0.034	-0.083	0.043	0.039	-	-	0.016	0.048 & 0.062	0.055
A1+N	A1, with sample-size adjusted variance	Year	-0.019	0.035	-0.085	0.035	0.035	-	-	0.016 (0.031)	0.052 (0.013)	0.055
A2+N	A2, with sample-size adjusted variance	Year	-0.019	0.035	-0.085	0.036	0.035	-	-	0.016 (0.031)	0.052 (0.013)	0.055
Disaggregated data												
D0_unnest	Island:Year:Month	Island:Year:Month	-0.033	0.032	-0.068	0.038	0.035	0.151	-	-	0.385	0.382
D1_unnest	Island:Year	Island:Year	-0.019	0.035	-0.085	0.035	0.035	-	0.055	-	0.408	0.407
D0	Year/Island/Month	Year/Island/Month	-0.032	0.034	-0.069	0.040	0.037	0.149	0.000	0.023	0.385	0.382
D1	Year/Island	Year/Island	-0.019	0.035	-0.085	0.035	0.035	-	0.052	0.016	0.408	0.407
D0+uneqvar	D0 with unequal variance for the islands	Year/Island/Month	-0.032	0.034	-0.069	0.040	0.037	0.149	0.000	0.023	0.381 & 0.388	0.382
D1+uneqvar	D1 with unequal variance for the islands	Year/Island	-0.019	0.035	-0.085	0.035	0.035	-	0.052	0.016	0.404 & 0.411	0.407

Table 7a: Repeat of Table 6a for **East Coast chick condition** data.

	Description	Random effects	St Croix delta se	Bird delta se	ave(se)	Random effects sd			Residuals sd	
						Year,Island,Month	Year,Island	Year	RES_sd1	RES_sd2
Aggregated data										
A1	Original model	Year	0.018 0.065	-0.029 0.067	0.066	-	-	0.076 (0.029)	0.082 (0.018)	0.070
A2	Unequal variance for the islands	Year	-0.005 0.088	-0.007 0.050	0.069	-	-	0.074	0.112 & 0.018	0.084
A1+N	A1, with sample-size adjusted variance	Year	0.026 0.064	-0.035 0.070	0.067	-	-	0.076 (0.028)	0.071 (0.020)	0.070
A2+N	A2, with sample-size adjusted variance	Year	0.026 0.064	-0.035 0.069	0.067	-	-	0.075 (0.027)	0.070 (0.020)	0.070
Disaggregated data										
D0_unnest	Island:Year:Month	Island:Year:Month	0.059 0.053	0.068 0.055	0.054	0.148	-	-	0.365	0.361
D1_unnest	Island:Year	Island:Year	0.031 0.066	0.009 0.068	0.067	-	0.103	-	0.379	0.378
D0	Year/Island/Month	Year/Island/Month	0.043 0.061	0.025 0.063	0.062	0.134	0.000	0.064	0.365	0.362
D1	Year/Island	Year/Island	0.026 0.063	-0.035 0.065	0.064	-	0.071	0.075	0.379	0.378
D0+uneqvar	D0 with unequal variance for the islands	Year/Island/Month	0.043 0.061	0.025 0.063	0.062	0.133	0.000	0.064	0.374 & 0.358	0.362
D1+uneqvar	D1 with unequal variance for the islands	Year/Island	0.026 0.063	-0.035 0.064	0.064	-	0.070	0.075	0.388 & 0.373	0.378

Table 8: Results for likelihood ratio tests which were conducted to evaluate whether nested or unnested models are to be preferred. For each case, two models (one nested, the other unnested) are compared. The log-likelihood values are listed (note that these are log-likelihood and not negative log-likelihood values, so the less negative value indicates preference), and yields the associated p-value for the comparison. “Neither” in the “Preferred model” column indicates that the nested model is not statistically better than the unnested model.

					InL			
		Model 1	Model 2		Model 1	Model 2	p-value	Preferred model
Max. Dist.	WC	D0_unnest	D0	BirdID included	-300.70	-288.61	5.64E-06	D0 (i.e. nested)
		D1_unnest	D1	BirdID excluded	-297.48	-296.74	0.22	Neither
	EC	D0_unnest	D0	BirdID included	-503.63	-491.65	6.24E-06	D0 (i.e. nested)
		D1_unnest	D1	BirdID excluded	-501.34	-498.65	0.02	D1 (i.e. nested)
Condition	WC	D0_unnest	D0	Month included	-5299.50	-5299.20	0.78	Neither
		D1_unnest	D1	Month excluded	-5762.10	-5762.00	0.80	Neither
	EC	D0_unnest	D0	Month included	-1617.00	-1614.90	0.12	Neither
		D1_unnest	D1	Month excluded	-1710.90	-1709.70	0.11	Neither

Appendix B: Exploring the relationship between chick condition and survival rate to evaluate what threshold for the island closure effect parameter δ corresponds a 1% change in annual penguin population growth rate

Note that this was originally the Appendix of FISHERIES/2021/APR/SWG-PEL/24; for reference purposes the same table and figure numbering has been retained.

In analyses of penguin data to date, estimates of the fishing effect parameter δ have been taken (on earlier advice by the Panel) to be biologically meaningful if $\delta < -0.1$ in cases where the response variable is analysed in log-space. This threshold was based on changes in survival estimates, and corresponds to a value below which (i.e. a δ value more negative than -0.1) analysis indicates that the population growth rate will decrease by 1% or more if fishing occurs in the neighbourhood of the island.

Details of the derivation of this threshold are provided in Robinson *et al.* (2014) (MARAM/IWS/DEC14/Peng/B4). The question this Appendix aims to answer is whether the same threshold of -0.1 is appropriate for other penguin response variables as has been assumed to date, specifically for chick condition. Chick condition data are investigated using the following approach: (1) establish a relationship between chick condition and survival, and (2) substitute this relationship in the equations of Robinson *et al.* (2014) to establish how the δ estimate for chick condition (where this is evaluated in normal- rather than in log-space, and consequently is now termed $\tilde{\delta}$) relates to a change in penguin population growth rate.

Figure A1 plots the predicted relationship between chick survival and chick condition for the Macaroni penguin population, extracted from Sherley *et al.* (2018). For this extraction, a web-based plot digitiser was used to extract a series of points from the solid black line of Figure S8(C) of Sherley *et al.* (2018) (which shows a predicted relationship between survival and condition) and a linear regression was implemented to retrieve the equation for this straight line. The resulting relationship is:

$$Survival = 0.2321 \cdot Condition + 0.1316 \quad (A1)$$

Figure A2 plots histograms of the South African chick condition data by coast, along with the estimates of survival that are derived from the chick condition data using equation (A1).

In order to calculate a $\tilde{\delta}$ threshold estimate for chick condition and the penguin population growth rate, essentially the same analysis is followed as in Appendix B of MARAM/IWS/DEC14/Peng/B4. For ease of reference, equations (B.5) and (B.6) from that paper are repeated below:

$$\eta^4 = \eta^3 S + H S^3 \quad (A2)$$

$$\Delta\eta = \frac{S^3}{4\eta^3 - 3\eta^2 S} \Delta H \quad (A3)$$

where η is the population annual proportional growth $\eta = N_{y+1}/N_y$, S is the mature female annual survival proportion and H is a measure related to the product of egg production and fledging success (hence encompassing chick survival).

Consider a steady state so that $\eta = 1$, then these equations become:

$$1 = S + H S^3 \quad (A4)$$

$$\Delta\eta = \frac{S^3}{4 - 3S} \Delta H \quad (A5)$$

Thus, for a “meaningful” change (decrease) in population growth rate of 1%:

$$\Delta H = \frac{\Delta\eta(4-3S)}{S^3} = (-0.01)(4-3S)/S^3 \quad (\text{A6})$$

Table A1 lists the mean chick condition values for the East Coast, West Coast and coast combined. The coast-combined mean chick condition value is $\overline{Condition} = 0.310$. From the chick survival-condition relationship (Equation A1), $\overline{Survival} = 0.204$, so that $H = 0.204$. Solving equation A4 then yields $S = 0.867$ (note that is very similar to the $S = 0.88$ evaluated by Robinson *et al.* 2014 from data for the Robben island population). Equation A5 consequently becomes:

$$\Delta\eta = 0.466 \Delta H \quad \text{or} \quad \Delta H = 2.144 \Delta\eta \quad (\text{A7})$$

For the (untransformed) chick condition response variable, the GLM equation used for the analysis is:

$$Condition_{y,i} = \alpha_y + \gamma_i + \tilde{\delta}_i X_{y,i} + \epsilon_{y,i} \quad (\text{A8})$$

where X is 1 if the island is open and 0 if it is closed. Thus $\tilde{\delta}$ is the change in chick condition resulting from opening to fishing compared to closing. Differentiating Equation (A8) w.r.t X gives $\Delta Condition = \tilde{\delta} \Delta X$. If X changes from zero to one (i.e. closed to open, so that the sign of $\Delta Condition$ is indicative of the impact of fishing, i.e. $\Delta Condition < 0$ implies a negative impact of fishing), then $\Delta X = 1$ and $\Delta Condition = \tilde{\delta}$.

Hence from the linear regression in Fig A1, $\Delta H = 0.2321 \tilde{\delta}$, and for $\Delta\eta = -0.01$ (i.e. a 1% decrease in the population growth rate resulting from closure, when $\eta = 1$) the critical (threshold) value of $\tilde{\delta}$ is:

$$\tilde{\delta} = \Delta H / 0.2321 = (2.144 \Delta\eta) / 0.2321 = 9.24 * (-0.01) \sim -0.092$$

The calculations above allow an estimate of δ (or $\tilde{\delta}$) to be converted to an estimate of the corresponding change in population growth rate, which is a more biologically intuitive quantity. The specific equations relating $\Delta\eta$ to δ and $\tilde{\delta}$ (taking $\eta = 1$ and $S = 0.867$) are:

$$\text{Maximum distance } (\delta^* \text{ in log-space}): \quad \Delta\eta = 0.094\delta \quad (\text{A9})$$

$$\text{Chick condition } (\delta^* \text{ in normal space}): \quad \Delta\eta = 0.108\tilde{\delta} \quad (\text{A10})$$

These equations have been used to provide the results shown in Figure 2.

The above calculations are for $\eta = 1$. If instead one assumes a decreasing population with $\eta = 0.95$, then $S = 0.819$, the $\tilde{\delta}$ threshold changes from -0.092 to -0.094, and equation (A10) changes to $\Delta\eta = 0.105\tilde{\delta}$, i.e. these threshold values are relatively insensitive to the value assumed for η .

Table A1: The mean chick condition values are reported in the first column, with the mean survival values estimated from Equation (A1) are listed in the second column. The third column lists the estimates of the mature female annual survival proportion S which follow when η is set to 1 in Equation (A2). Distributions of the chick condition data, as well as the survival rates estimated by the Equation (A1) relationship are shown in Figure A2.

	$\overline{Condition}$	$\overline{Survival} = m * \overline{Condition} + C$	S (from Equation A2 with $\eta = 1$)
(1) West Coast	0.327	0.208	0.865
(2) East Coast	0.259	0.192	0.872
(3) Coasts combined	0.310	0.204	0.867

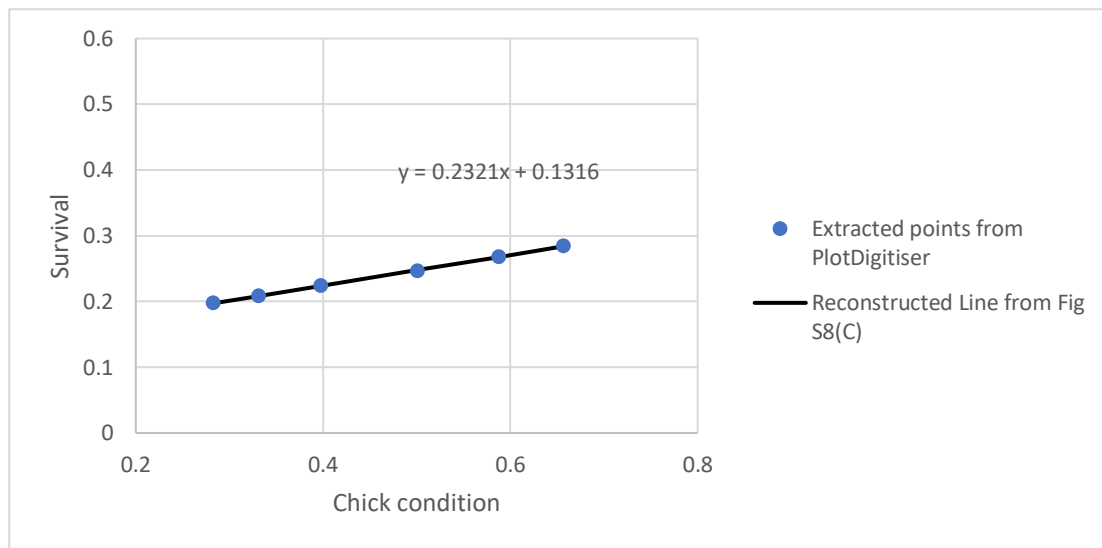


Figure A1: Extraction of the predicted relationship between chick survival and condition from Figure S8(C) of Sherley *et al.* (2018) for the Macaroni penguin population. The blue dots show the points extracted from the pdf, and the black line shows the regression line through these points, which is essentially a reconstruction of the solid black line of Figure S8(C) of Sherley *et al.* (2018).

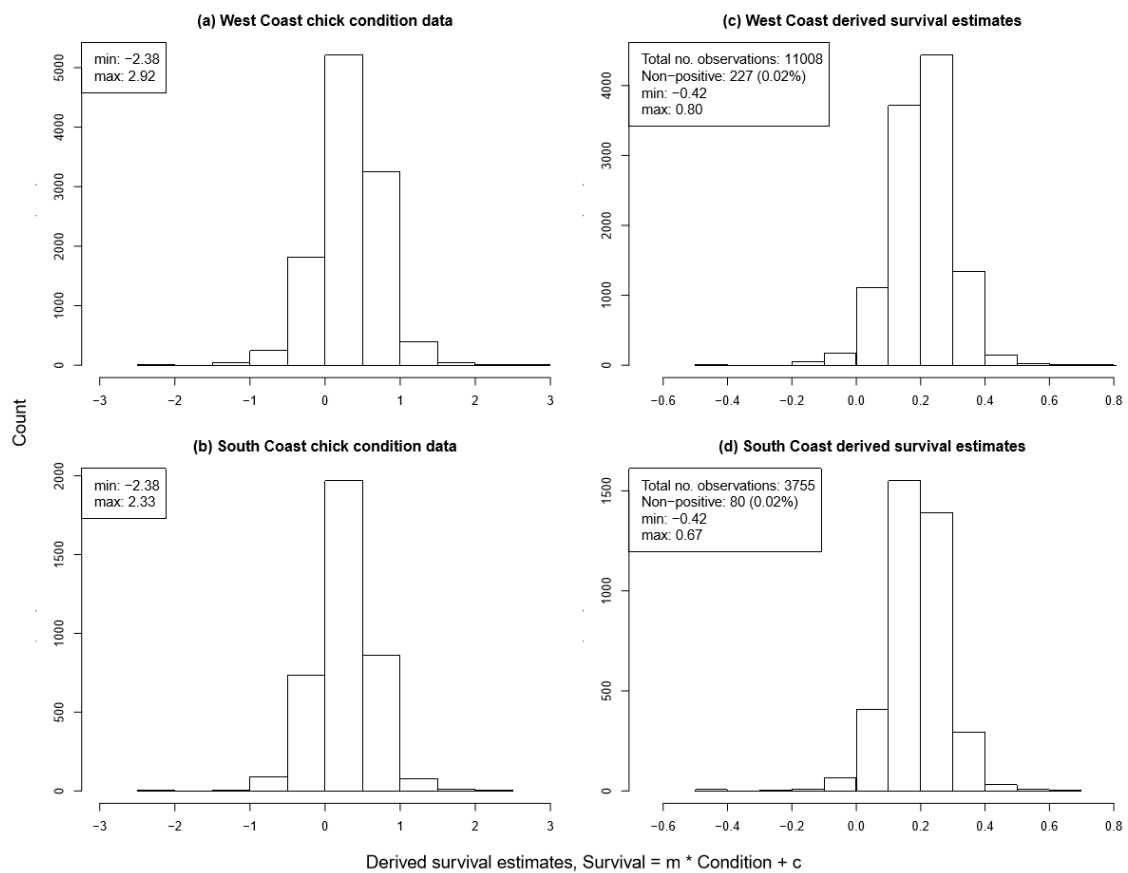


Figure A2: Histograms of the WC and EC chick condition data (left column) and the corresponding survival estimates obtained using the relationship in Figure A1.

Appendix C: Additional output from the month standardisation exercise.

This Appendix presents some additional output for the month standardisation exercise conducted. Figure C1 plots the estimated month effects for the various response variables and islands. Figure C2 plots the original aggregated series and those aggregated series derived from the standardised data. Figure C3 plots the residuals.

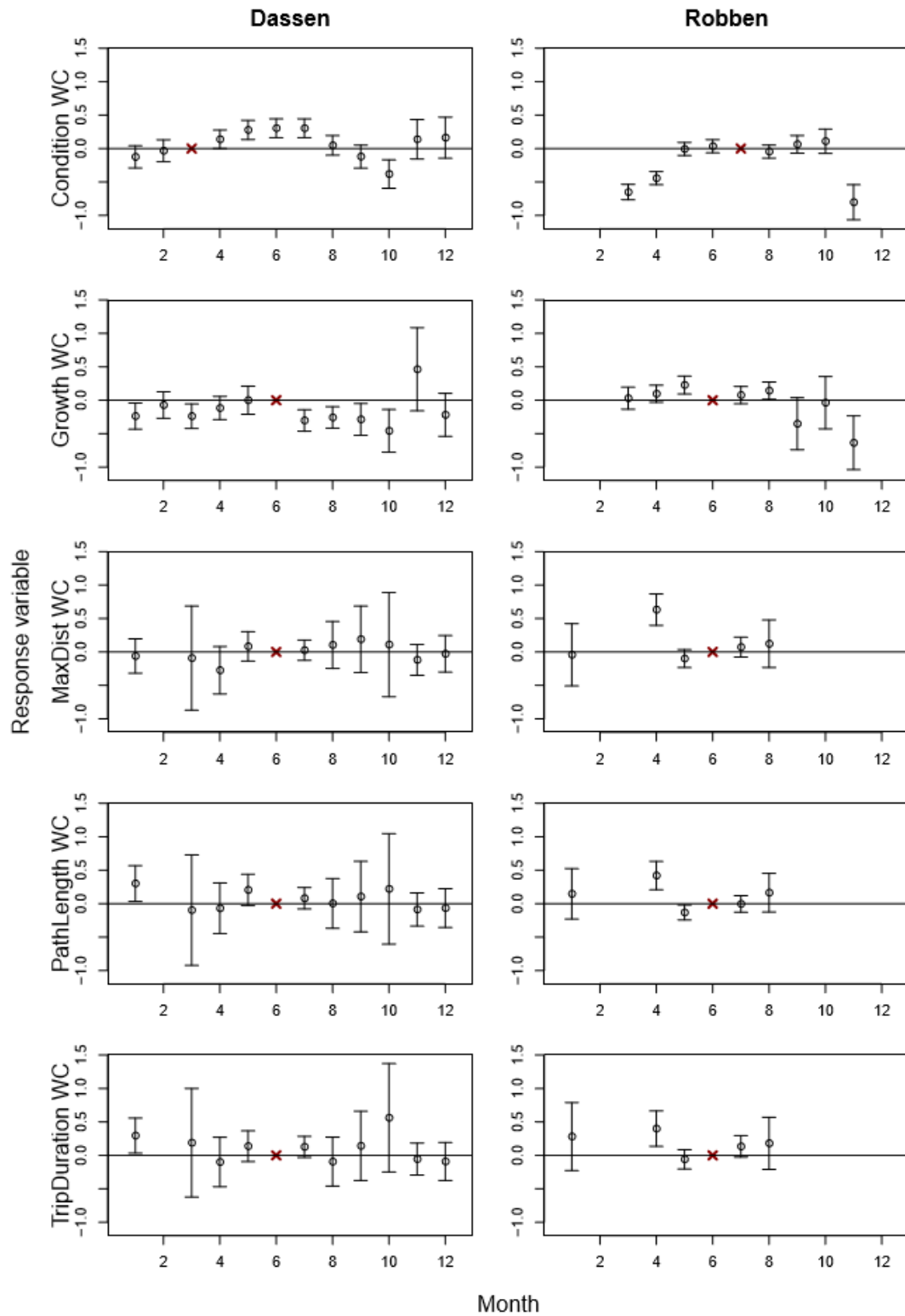


Figure C1a: Month effects estimated for the five response variables for the **WC islands**. Note that for comparison purposes the chick condition and chick growth effects (not analysed in log-space on account of negatives in the data set) have been standardised by dividing by the mean of the data points for the island in question so that all are expressed in relative terms and can be compared. For each plot, a red cross marks the month with the most data points, which was used as the reference level.

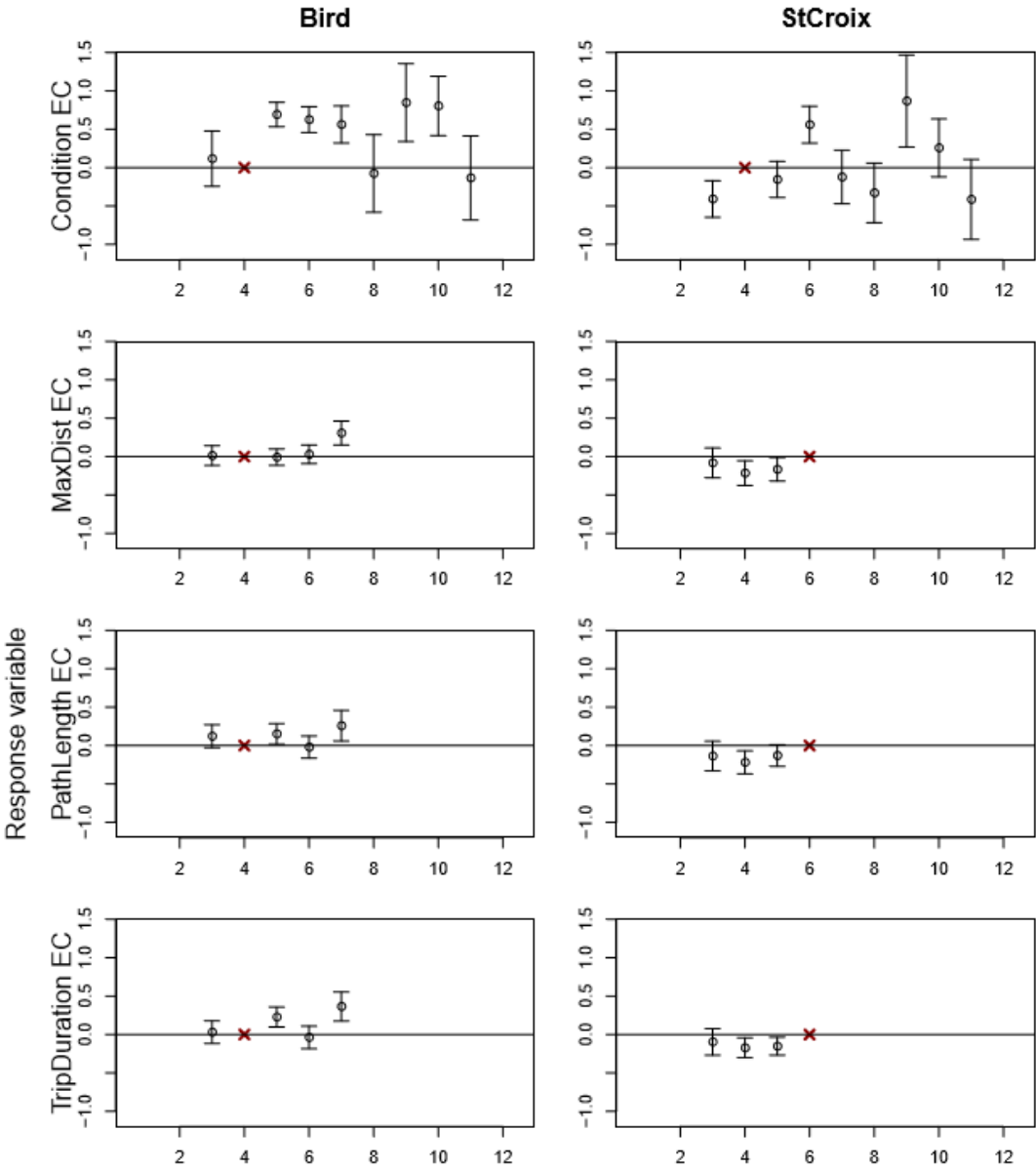


Figure C1b: Month effects estimated for the four response variables for the EC islands.

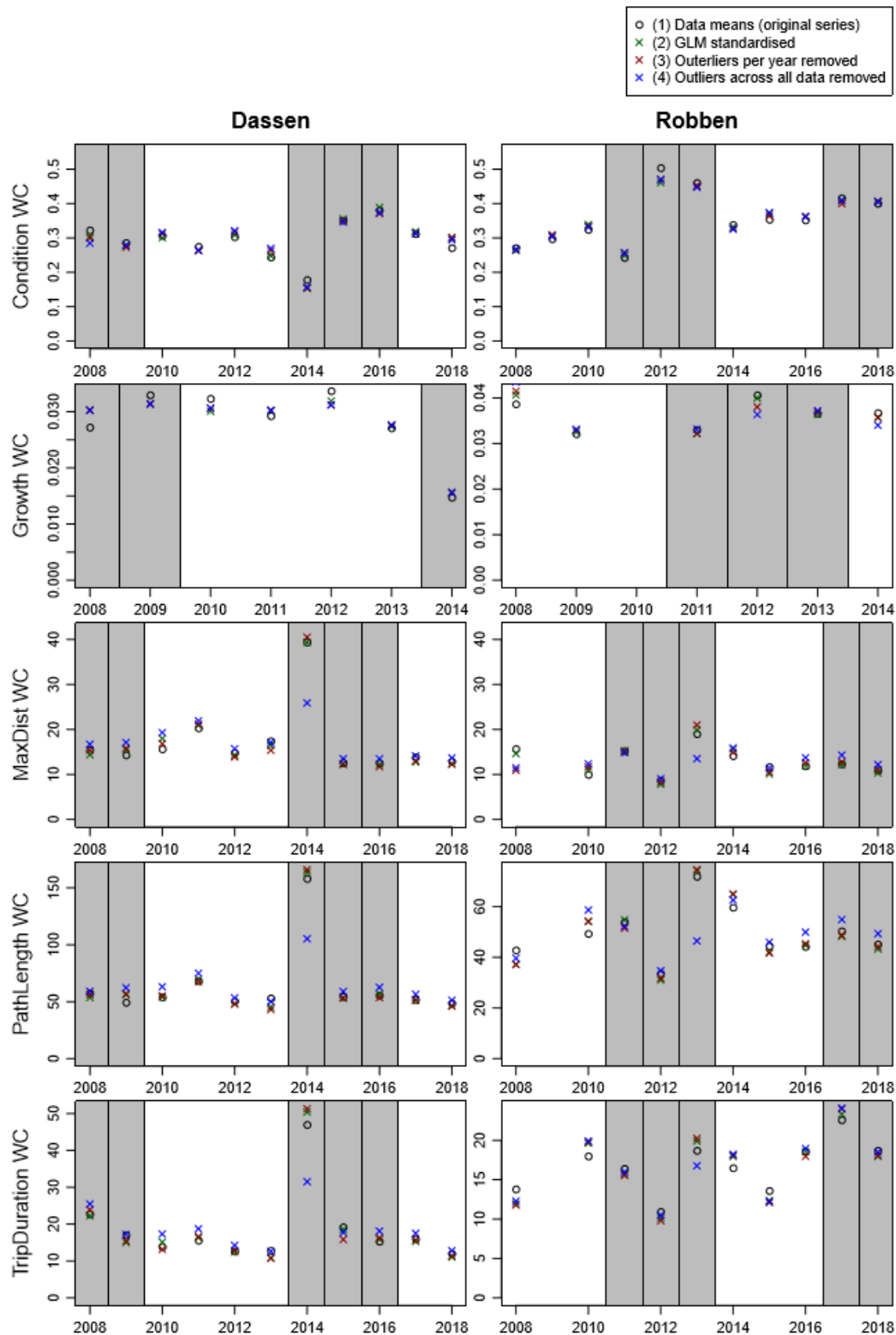


Figure C2a: Original and standardised aggregated series for the **WC response variables**. The first shows the original series, the second the standardised series with no outliers removed, the third with outliers identified according to each island/year pair and removed before aggregating (M1) and the fourth with outliers identified across all years for each island and removed (M2). The grey shaded areas indicate island closure in that year.

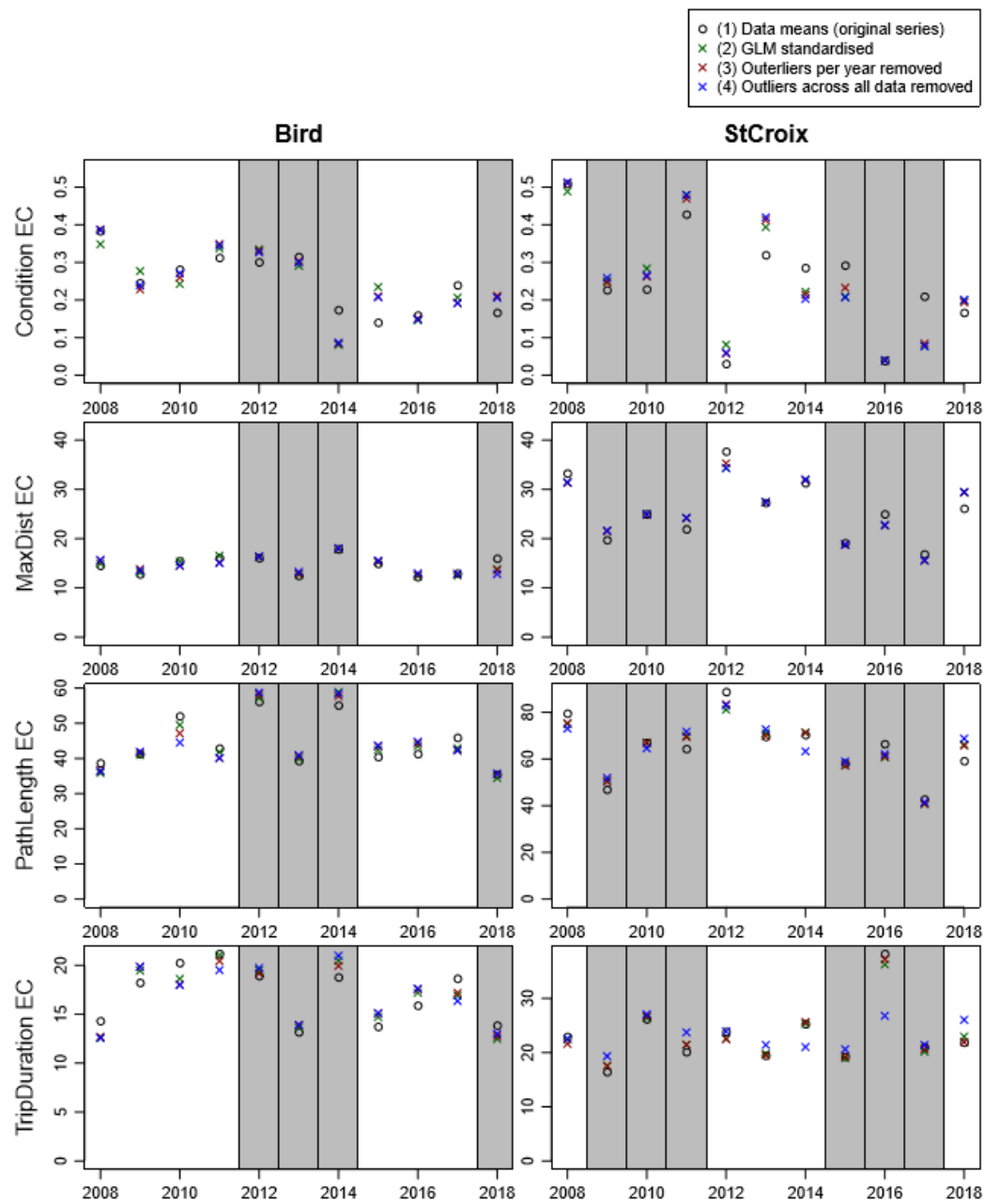


Figure C2b: Original and standardised aggregated series for the EC response variables.

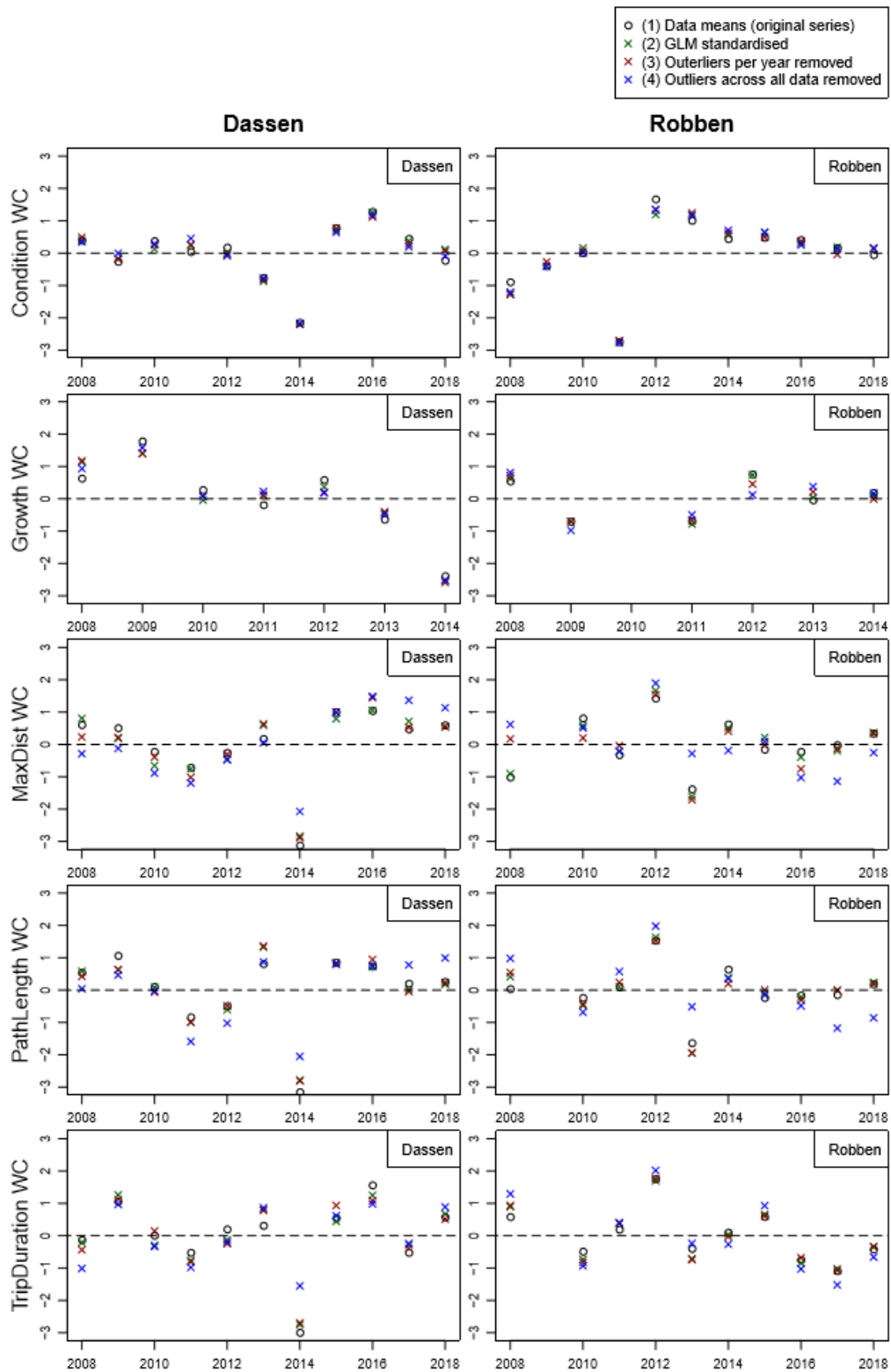


Figure C3a: Standardised residuals for the **West Coast aggregated** series shown in Figure C2a.

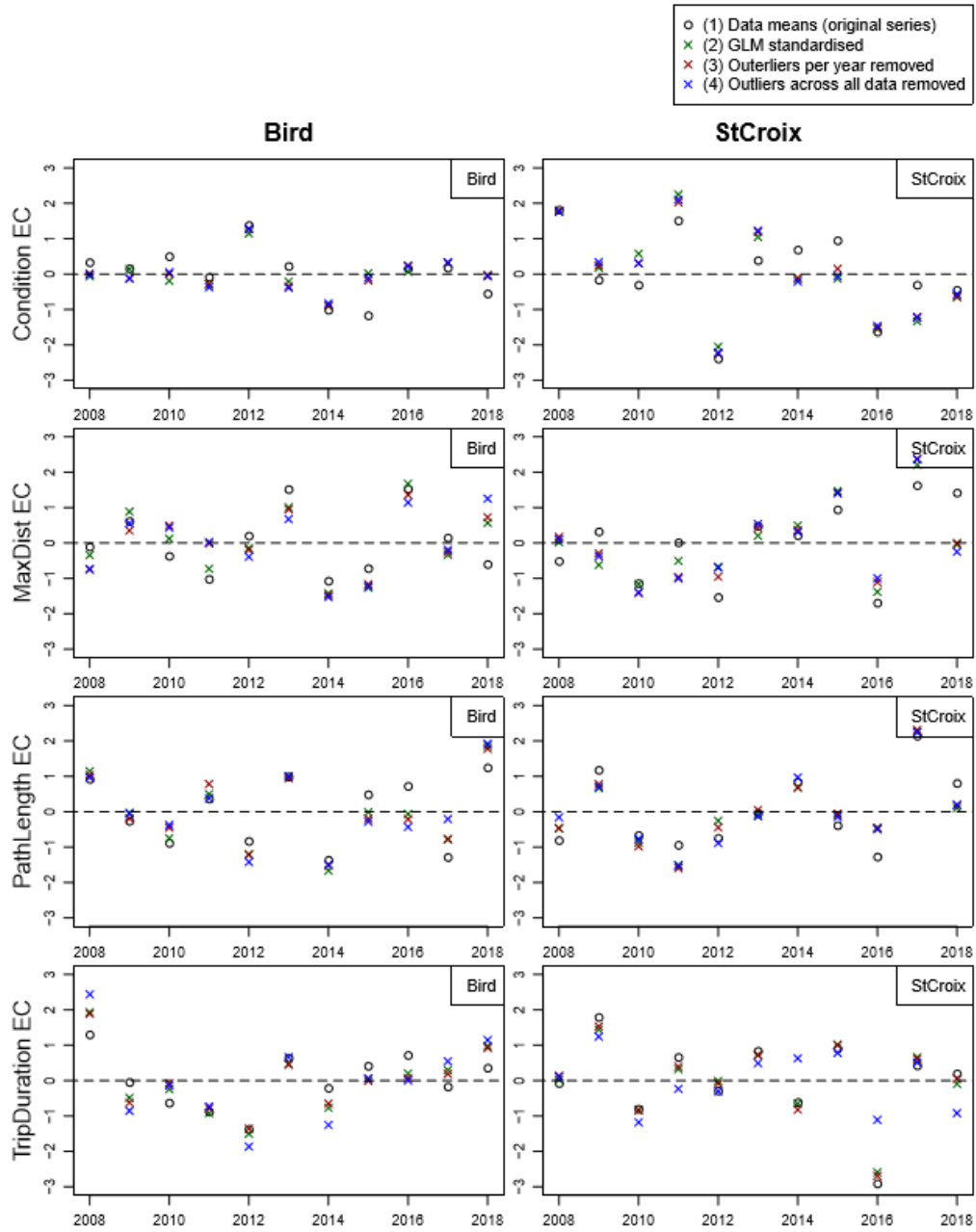


Figure C3b: Standardised residuals for the **East Coast aggregated** series shown in Figure C2b.