Application of the power analysis algorithm to chick condition data standardised by month

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

Contact email: *mlland028@myuct.ac.za*

Summary

The effect of a month on the estimation of the impact of fishing on penguin chick condition at Robben and Dassen islands is taken into account by GLM standardising the chick condition data for this co-variate. The difference to results is small, with comparable estimates of the impact moving slightly in the direction of a lesser impact of fishing compared to when the month co-variate is ignored.

Introduction

At the January PWG meeting, a request was made to take month into account in the power analysis procedure for the island closure experiment related to penguins, and chick condition as well as foraging trip data were proposed as candidates for this exercise. Following a closer examination of the foraging data, it became evident that these are not suitable for such an analysis, as there is only seasonal information (summer and winter) available in the data provided (to Janet Coetzee of DAFF), and furthermore data only for summer for Robben Island. In these circumstances, this exercise has been conducted for chick condition data only.

The approach followed has been to apply a simple additive GLM to the individual chick condition data points (i.e. not aggregated for each month), with year 2009 and month 6 assumed for the reference selections (note that these choices do not matter, as analyses were conducted on the logs of the annual factors). The equation for the standardisation is:

$$F_{y,m,i} = \alpha_y + \beta_m + \epsilon_{y,m,i} \qquad (1)$$

where $F_{y,m,i}$ is the chick condition response for year y, month m and chick i, α_y is the year effect, β_m is the month effect and $\epsilon_{y,m,i}$ is an error term. The model is applied to the data from each island separately, and standardised annual chick condition indices for year y are then given by α_y, for each island separately. An additional exercise was conducted where outliers were removed, corresponding to any individual data point for which the residual $\epsilon_{y,m,i}$ following the first standardisation differed by more than 3 standard deviations from zero.

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

Results

Table 1 summarises the power analysis results applied to four variants of the chick condition data. The first is original chick condition series, and the second where the 2004 data point for Robben Island has been removed from this series. This data point was included in the original data series and power analyses, but the data provided which have been used for the standardisation exercise here have individual observations data from 2008 onwards only, and because of this a series with the 2004 data point removed has been included to allow for direct comparison with the standardised series. The last two series are the chick condition standardised for month, and the standardised series when outliers are removed.

Figure 1 plots the original annual chick condition time series alongside the standardised series, and the standardised series with outliers removed. Figure 2 plots the GLM-bias adjusted estimates for the closure effect δ for the above-mentioned series, as well as for the additional series where the 2004 data point for Robben Island has been removed. Figure 3 shows comparison plots of the point estimates and 95% confidence intervals for the month effects estimated by the GLM for the two standardised series. Figure 4 plots the standardised residuals when the closure EM is applied to the data. Figure 5 plots the integrated detection probabilities against year given future simulated data.

Discussion

Standardising for month does not change the overall conclusions with regards to a biologically meaningful impact of fishing. The effect on Robben Island remains biologically meaningful - in fact the X value increases, but this arises primarily because of omission of the 2004 value. For Dassen Island, estimation of the effect of fishing remains inconclusive, and standardising for month has the effect of increasing (i.e. making less negative) the GLM-bias adjusted estimates of δ.

These results are not entirely unexpected, as the standardised series in Figure 1 do not show a distinct trend that would likely lead to a much more conclusive indication of the impact of fishing, such as a marked decrease in chick condition in open years and/or an improved chick condition in closed years.

Acknowledgements

Computations were performed using facilities provided by the University of Cape Town's ICTS High Performance Computing team: http://hpc.uct.ac.za. The data providers (DEA/CapeNature, Richard Sherley) are gratefully acknowledged.

Table 1: Summary of the power analysis results for the four chick condition series considered. Results are shown for the closure-only EMs, integrated over the five OMs. The closure-only OMs receive twice the weighting of the others to take into account that the catch-only and catch+closure OMs are implemented for two catch-biomass correlation values. The first column of numbers shows the GLM-bias-corrected estimates of δ and the second column the EM estimates of the standard error. The third column shows the area (corresponding to probability) under the normal curve with mean δ_{data}^{EM*} and standard deviation *se* that lies to the left of the Threshold (i.e. "X" from the main text). The fourth column lists the P_{min} values. If $X > P_{min}$ then there is evidence in the data of a fishing effect that is biologically meaningful with respect to penguin demographics. In such cases the table entry has been highlighted in grey. The next five columns show the values below which the true δ is not likely to be, for a range of (one-tailed) risk levels from which one might be adopted for decision purposes. Cells that have been highlighted in yellow indicate that the value below which δ is not likely to lie, is above the Threshold. The last column shows the number of years that the experiment would need to be continued before it is likely to be possible (with 80% probability) to conclude from the data that δ is less than the Threshold, IF the true δ is indeed below the Threshold.

		EM applied to data				δ_{crit} for a range of risk levels					
Island	Data type	δ_{data}^{EM*}	se	X	P_{min}	2.50%	5%	10%	20%	50%	Years
Dassen	Chick condition	-0.03	0.14	0.32	0.58	-0.24	-0.20	-0.16	-0.09	-0.03	>20
	Chick condition, excl 2004	-0.03	0.13	0.30	0.60	-0.22	-0.18	-0.14	-0.09	-0.03	>20
	Chick condition, standardised	0.01	0.11	0.18	0.63	-0.14	-0.12	-0.09	-0.06	0.01	>20
	Chick condition, remove outliers	0.04	0.10	0.09	0.64	-0.11	-0.08	-0.05	-0.03	0.04	>20
Robben	Chick condition	-0.14	0.13	0.62	0.61	-0.37	-0.31	-0.26	-0.16	-0.09	10[2]
	Chick condition, excl 2004	-0.20	0.13	0.78	0.63	-0.45	-0.39	-0.32	-0.19	-0.13	0
	Chick condition, standardised	-0.16	0.11	0.71	0.63	-0.36	-0.29	-0.25	-0.17	-0.11	2-5
	Chick condition, remove outliers	-0.18	0.10	0.78	0.65	-0.37	-0.32	-0.27	-0.19	-0.13	2-5

[2] It might seem counter-intuitive that for Robben Island chick condition, this estimate of how long the experiment needs to continue is non-zero, even though the conclusion is drawn that there is evidence to support a biologically meaningful impact of fishing on the penguin population. The values reported here (in the "Years" column) estimate how long the experiment would need to continue before it becomes likely (with 80% probability) that a fishing effect would be detected if it is there. However, a detection probability of 40-50% does not preclude the possibility of already having detected a meaningful effect now, rather a higher likelihood of having missed detecting it.

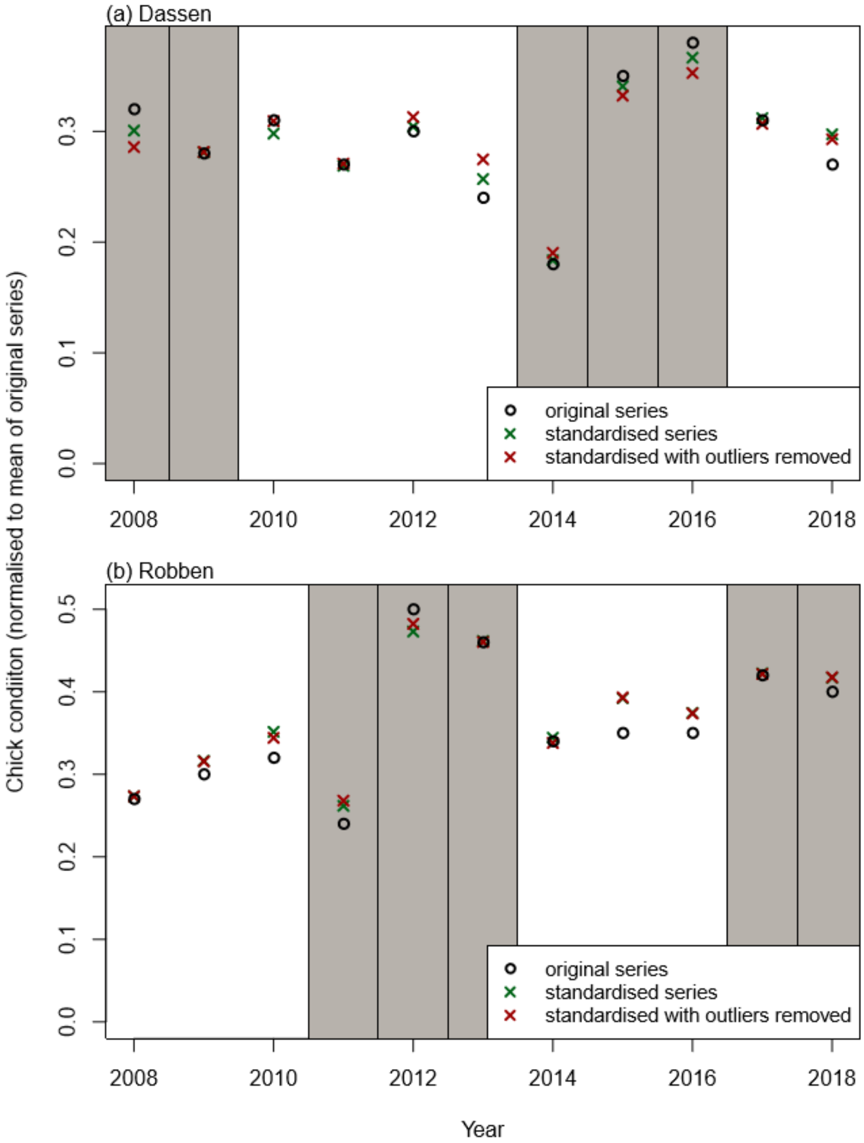

Figure 1: Original chick condition series, series standardised by month, and series standardised by month after removing outliers (corresponding to points with residuals more than three standard deviations from the mean for the first standardisation). The two standardised series have been normalised to have the same mean as the original series (note that this does not affect the further analyses, which are based on the logs of these values). Grey shading has been used to indicate years in which each island was closed to the fishery.

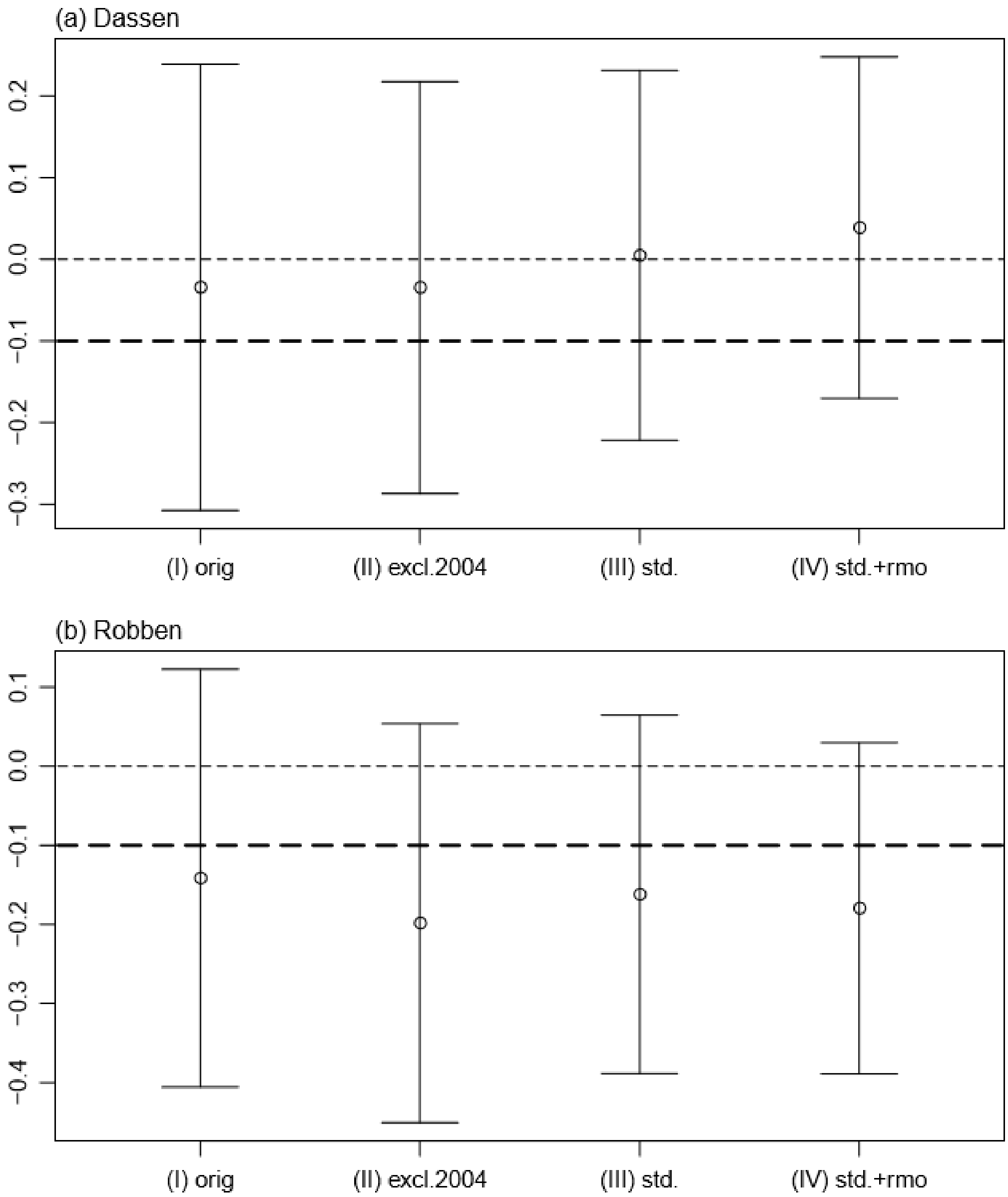

Figure 2: GLM-bias adjusted estimates of δ are shown along with the associated rough 95% confidence interval (corresponding to twice the standard deviation from the mean) for (I) the original chick condition series, (II) the original chick condition series with the 2004 data point for Robben Island removed (III) the chick condition standardised for month and (IV) chick condition standardised for month, after outliers had been removed (where an outlier is defined as any data point for which the residual after the first fit (III) is more than three standard deviations away from the standardised mean).

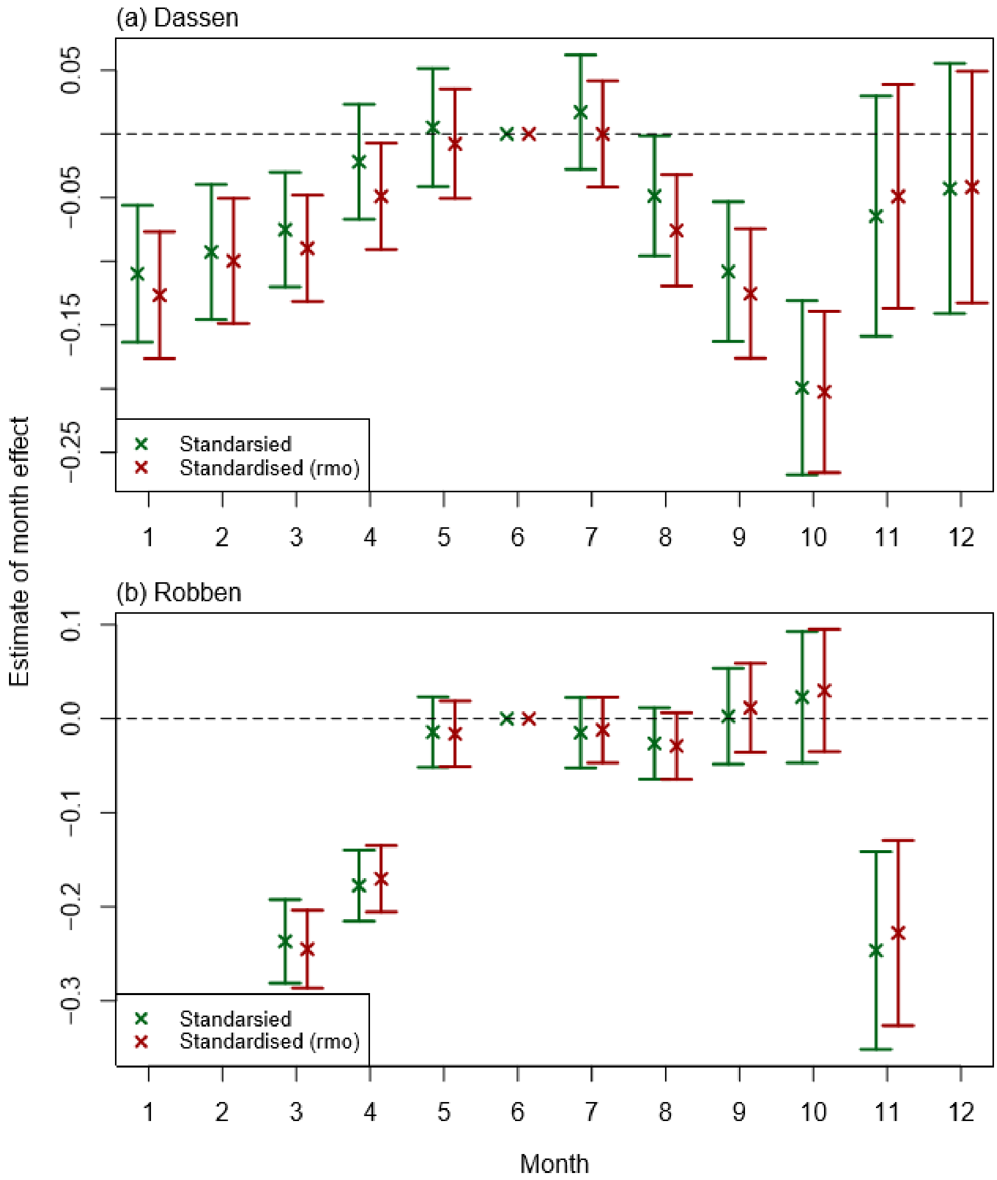

Figure 3: Plots of the month effect estimates from the GLM fits for the chick condition standardised for month (green points) and standardised for month after outliers have been removed. Point estimates are shown with rough 95% confidence intervals corresponding to two standard deviations from the point estimate. Note that there is no confidence interval shown for month 6 as this was selected as the reference month.

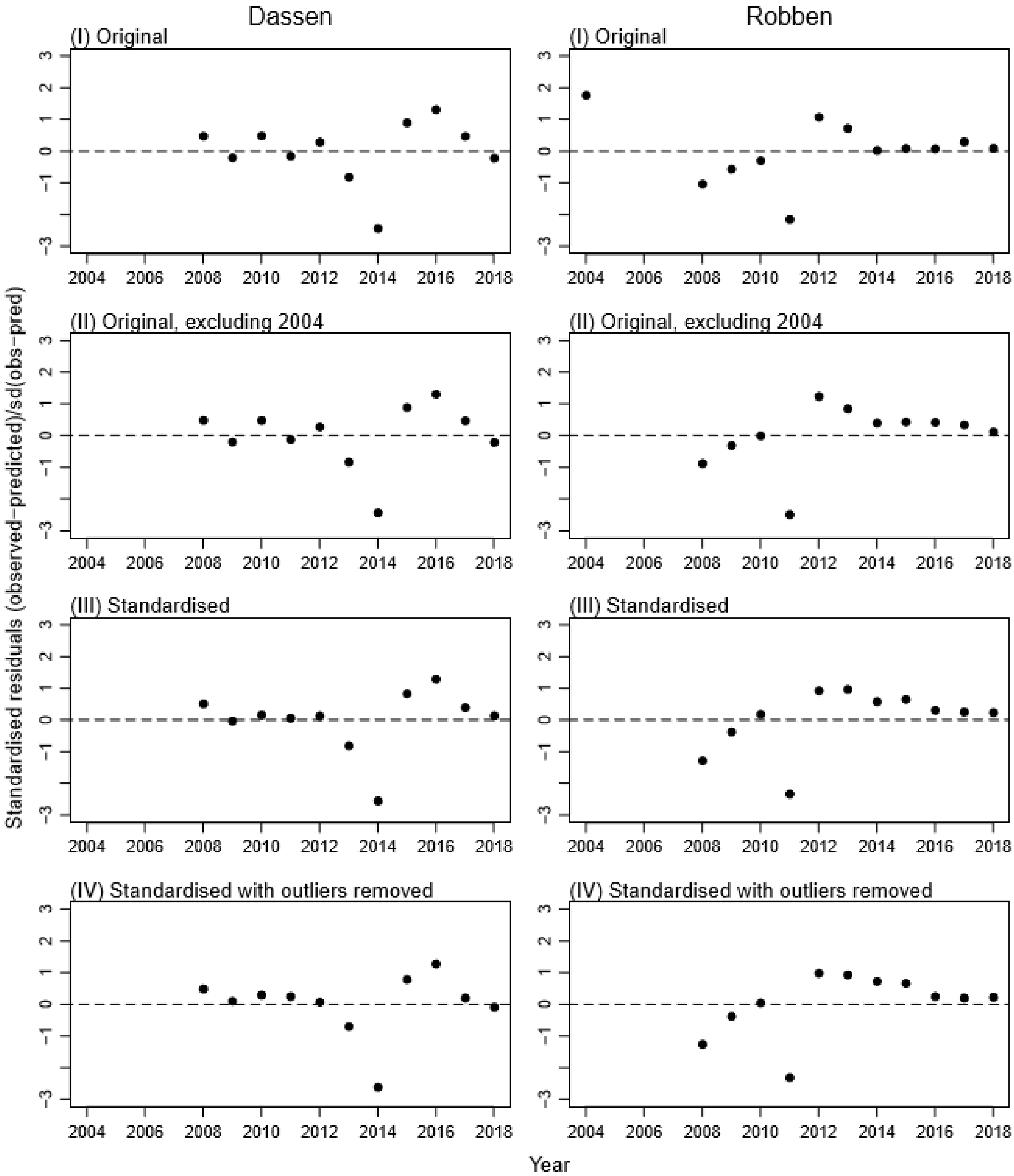

Figure 4: Plots of the standardised residuals when the closure only EM is applied to the four series considered.

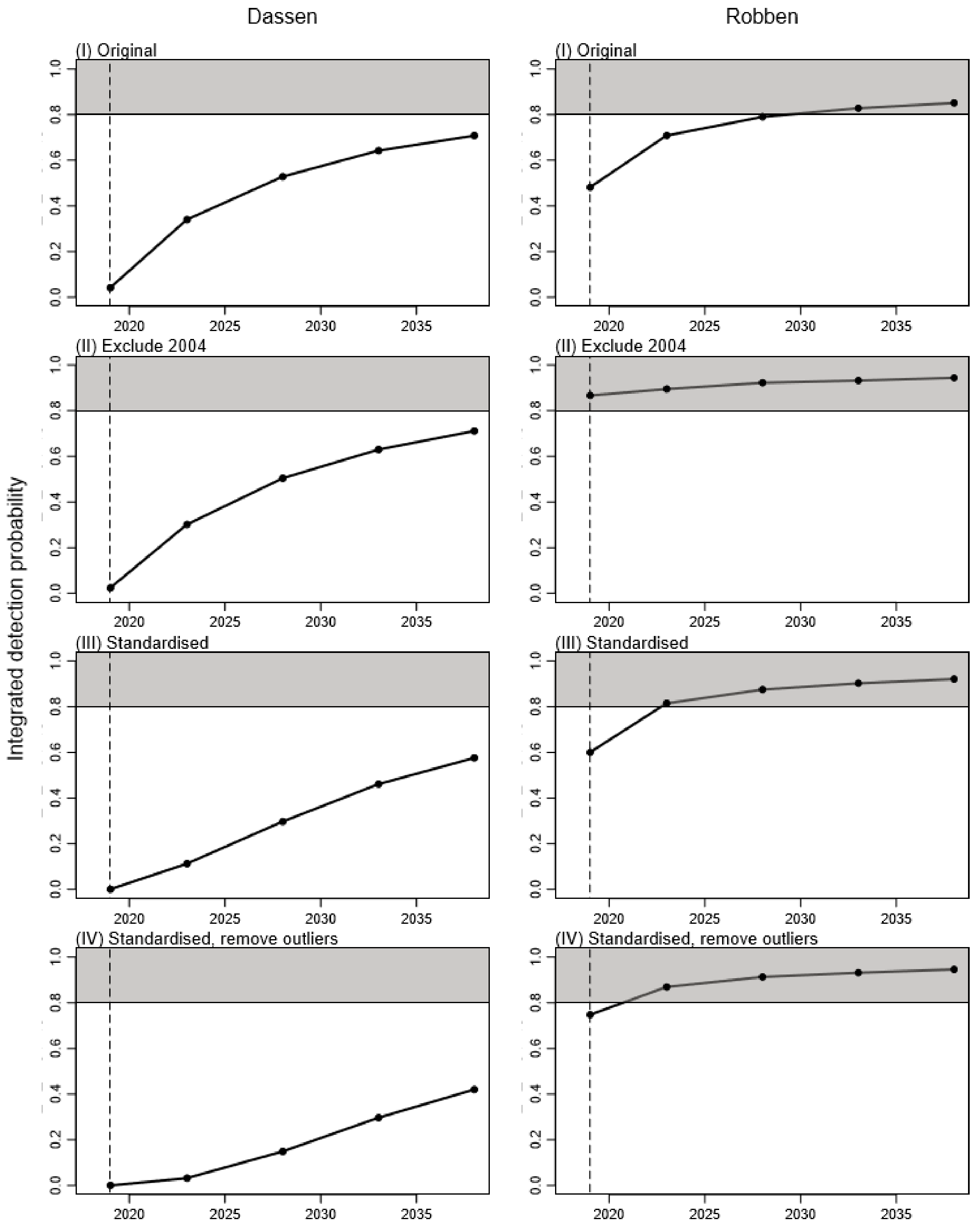

Figure 5: Plots of the integrated detection probabilities as projected over time. The probabilities have been weighted across OMs to produce a single detection probability curve for each EM, and show the length of time the closure experiment would need to be continued for before a biologically meaningful fishing effect is likely to be detected from the data, if such an effect is present. Note that this is the time likely needed to reach an 80% probability of detecting an effect if it is there, and does not preclude the possibility of detecting such an effect earlier – there is just a higher probability of missing it.