A regression-based estimate of the November 2020 sardine recruitment strength based on the June 2021 recruitment survey result

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

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

Summary

The regression method developed in 2020 to predict the November 2019 sardine recruitment strength from the estimate from the 2020 recruitment survey is shown to have performed very well (given that the assessment-based estimate of that recruitment is now available). Applying the same method to the result from the 2021 recruitment survey predicts a low November 2020 sardine recruitment compared to 2019, with a 90% probability of being between 8-40% of the November 2019 level.

Key words: Sardine, recruitment, regression, survey, assessment

Introduction

For the sardine population, the November recruitment in year $y-1$ is sampled by the recruitment survey in year y. The nature of the stock assessment is such that the recruitment survey result in year y is available prior to the assessment update which provides an estimate for the November recruitment in year $y-1$. By assuming that the historical recruitment survey results are directly proportional to the assessment estimate of recruitment for November of the previous year, an attempt can be made to predict the November recruitment of year $y-1$ based on the survey result of year y. The assumed relationship is:

$$ln\, surv(y) = ln\, R(y) + ln\, k + eps(y) \qquad\qquad eps(y) \sim N(0, sig2) \qquad (1)$$

where $R(y)$ is the historical assessment estimate of November recruitment in calendar year $y-1$ and $surv(y)$ is the corresponding recruitment survey result for historical year y. Note that for this document, years shown in the plots correspond to the year of the recruitment survey.

A similar but more complicated exercise was undertaken last year (Butterworth and Ross-Gillespie 2020). Since then, the assessment November recruitment series has been updated (de Moor 2021). Accordingly, two main analyses are conducted here:

(I) Regress the updated November recruitment series (values up to Nov 2018) against the survey series (values up to autumn 2019) to see what the prediction for the November 2019 recruitment estimate would have been, and how that compares to the actual November 2019 estimate from the assessment update.

(II) Regress the entire updated November recruitment series (values up to Nov 2019) against the entire survey series (values up to autumn 2020) to predict what the estimate for the November 2020 recruitment will be, together with the associated uncertainty.

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

Results

Results are shown for three regression options: (a) the entire series, (b) values from 2005 onwards only and (c) the 2005 onwards but with the 2010 point excluded. The analyses were run for (b) and (c) to be consistent with approach taken last year. A plot of *surv(y)* against *R(y)* is shown for the three options in Figure 1, along with the fit from equation (1). A plot of the residuals *eps(y)* against year from the regression fit is shown in Figure 2 for all options, where $eps(y) = ln\ surv(y)\text{-}\ ln\ surv\_hat(y)$. The estimate of *sig* for option (a) is 0.45, for option (b) 0.48 and for option (c) 0.49.

Table 1 lists the data used in the analyses. Figure 3 provides further diagnostic plots for the fits of equation (1) to the data for option (a) and (b). Figure 4 plots log-normal distributions for analyses (I) and (II) where the median of the distribution is (I) the expected 2019 November recruitment estimate that would have been predicted last year using the current assessment recruitment series (plotted alongside the actual 2019 estimate that is now available) and (II) the expected 2020 November recruitment estimate based on the 2021 survey result. The CVs of the normal distributions are given by the standard deviation of the *epsilon* values of the regression analysis. Figure 5 plots of the November recruitment estimates against year to compare with the point estimates from the regressions.

Discussion

It is immediately evident from the top panel of Figure 4 that all three regressions from 2020 provided very good estimates of the final November 2019 recruitment of 42.7 as determined from the subsequent assessment (which made use also of the November 2020 survey result).

Repeating this estimation process one year later using the 2021 sardine recruitment survey estimate provides November 2020 sardine recruitment estimates ranging from 7.0 to 8.3, with a lower 5th%ile from 3.19 to 3.98 and upper 95th% from 15.21 to 17.24.

Very clearly, this approach indicates a November 2020 sardine recruitment to be much less than that the year before – at a fraction between 0.08 and 0.40 thereof. The point estimates are close to the second lowest historical value.

References

Butterworth, D.S, and Ross-Gillespie, A. 2020. Continuous approach applied to sardine recruitment survey estimate to advise on juvenile sardine TAB estimates for 2020 and the corresponding recruitment strength. DEFF Fisheries document FISHERIES/2020/JUL/SWG-PEL/60.

de Moor, C.L. 2021. Updated assessment of the South African sardine resource using data from 1984-2020. DEFF Fisheries document FISHERIES/2021/APR/SWG-PEL/23.

Table 1: Data used in the analyses. The assessment estimates for November recruitment (in billions of fish) in year $y-1$ are linked to (and regressed against) the recruitment survey estimate in year y, which is why two year columns are listed. The regression for option (a) uses the data for all years to 2020, option (b) excludes data prior to 2005 and option (c) further excludes the year 2010. Note that throughout this document, where references are made to years such as in the preceding sentence, these values correspond to the year of the recruitment survey estimate. The assessment estimates of November recruitment were provided by C.L de Moor (as per analyses of de Moor 2021), and the recruitment survey estimates provided by J. Coetzee (*pers. comm.*). Note that the 2020 recruitment survey estimate of 11.777 was updated after Butterworth and Ross-Gillespie (2020) was tabled, and hence differs from the value of 7.01 used in that document.

Assessment Year	Assessment estimate of November recruitment	Survey Year	Recruitment survey estimate
1984	10.3816	1985	3.592
1985	15.3525	1986	3.691
1986	20.1740	1987	7.380
1987	5.4273	1988	0.440
1988	13.9012	1989	2.137
1989	16.1823	1990	2.502
1990	11.0028	1991	1.915
1991	25.8954	1992	5.633
1992	64.1250	1993	15.238
1993	9.0961	1994	2.654
1994	60.2353	1995	25.388
1995	13.3146	1996	3.204
1996	113.0330	1997	36.856
1997	61.3317	1998	10.716
1998	65.4672	1999	10.378
1999	81.3913	2000	20.002
2000	205.8350	2001	60.065
2001	173.7780	2002	49.153
2002	147.8270	2003	36.448
2003	22.3562	2004	4.089
2004	12.5670	2005	2.858
2005	21.6173	2006	9.506
2006	11.0874	2007	2.995
2007	18.0642	2008	4.090
2008	20.2347	2009	9.289
2009	84.3697	2010	35.569
2010	16.2999	2011	5.799
2011	25.4825	2012	7.986
2012	27.7333	2013	12.586
2013	8.9938	2014	1.985
2014	10.4432	2015	6.258
2015	6.7885	2016	0.811
2016	12.8972	2017	7.180
2017	17.6661	2018	-
2018	24.7320	2019	3.540
2019	42.7407	2020	11.777
2020		2021	2.140

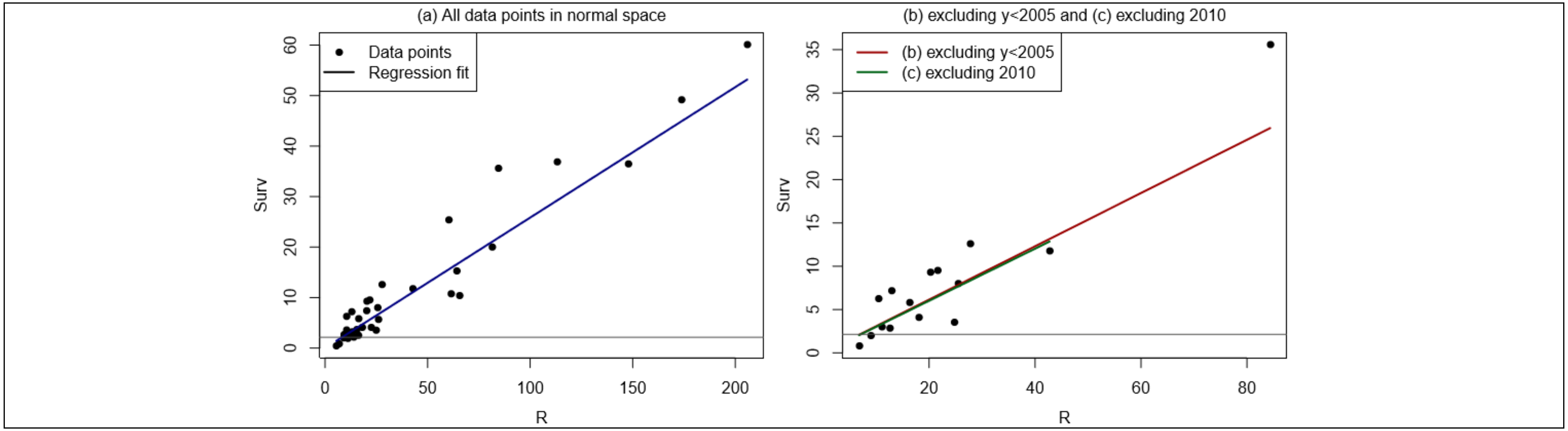

Figure 1: Plots of the data points in normal space for (a) 1985-2020, (b) 2005-2020 and (c) 2005-2020, but excluding 2010. The 2021 survey result of 2.14 is indicated by the grey horizontal line.

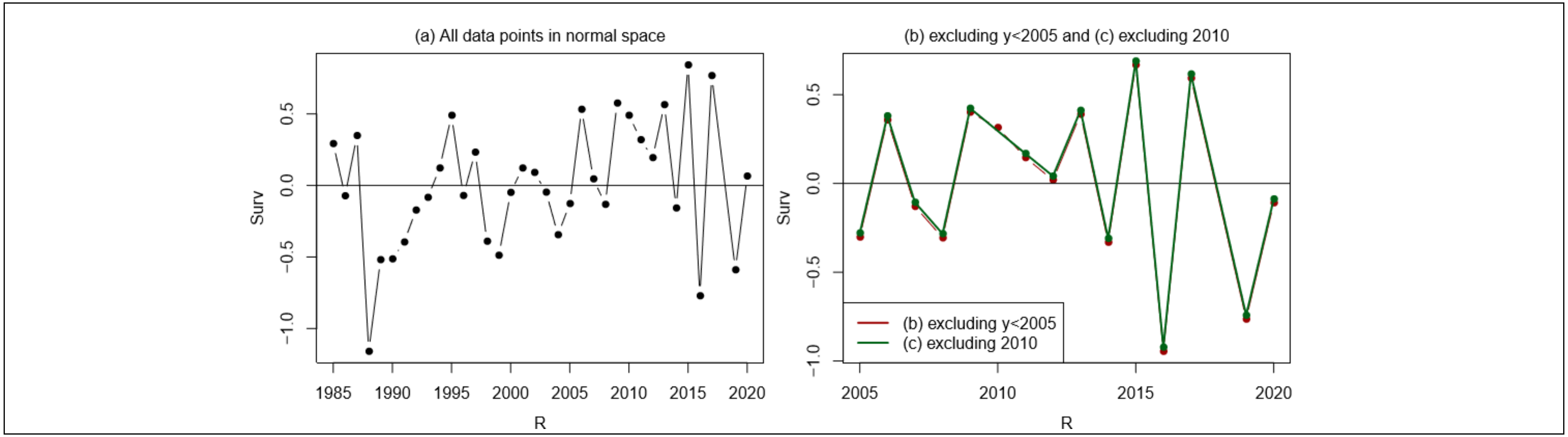

Figure 2: Plots of the residuals of the fit of equation (1) against year for Analysis (II) and (a) 1985-2020, (b) 2005-2020 and (c) 2005-2020, but excluding 2010.

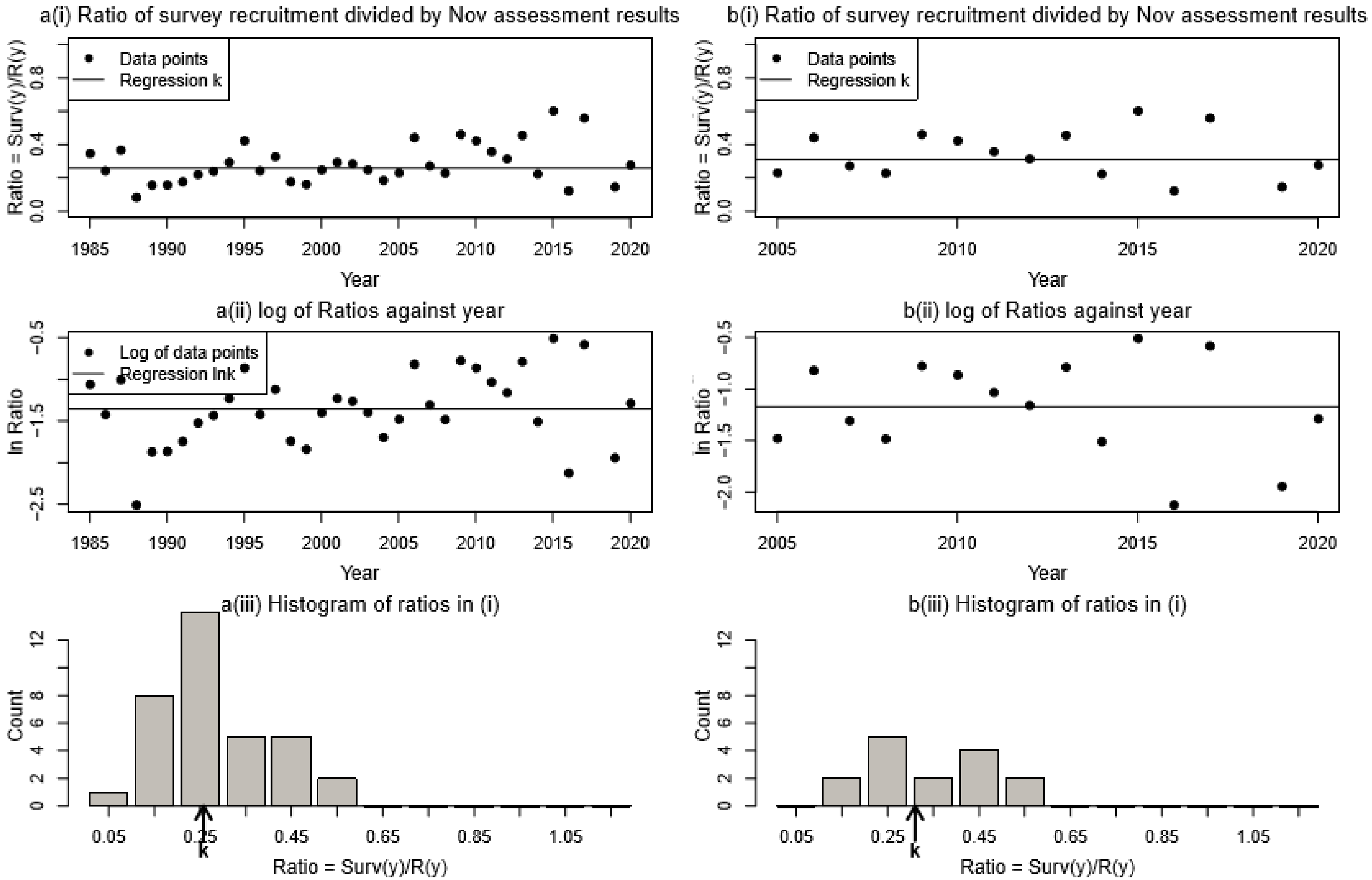

Figure 3: Further diagnostic plots related to the data used for the fits of equation (1), showing those for 1985 to 2020 on the left and for 2005 to 2020 on the right. Plot (ii) in terms of the logarithms of the ratios is added with a view to greater symmetry in a situation where the values in (i) are necessarily positive. The estimates of *k* from equation (1) are indicated.

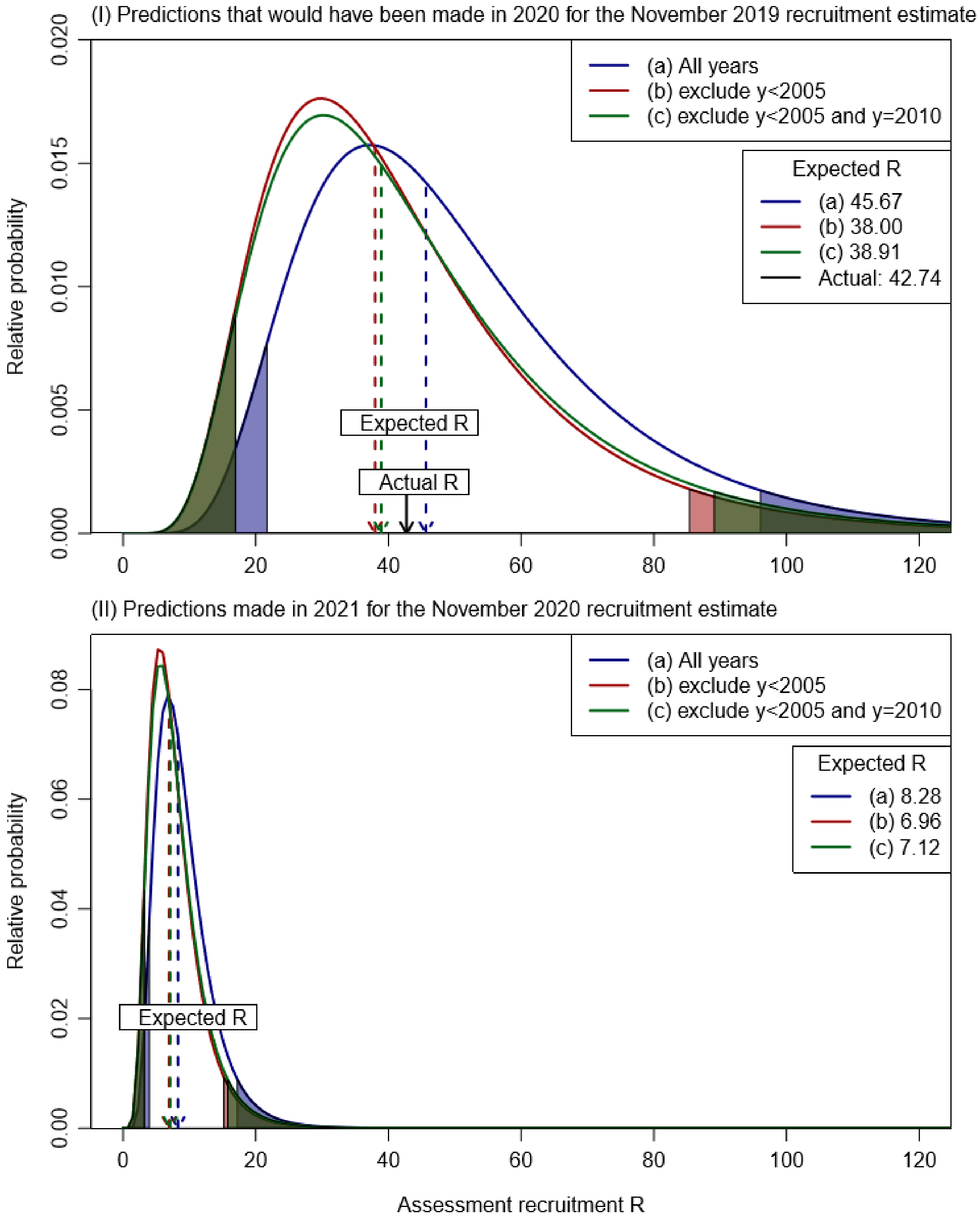

Figure 4: Probability density function plots for the log-normal distribution for the expected November recruitment given the regression results. The top plot (a) shows the predictions that would have been made in 2020 given the current recruitment series at that time (i.e. the current series truncated at 2019) along with the actual November 2020 recruitment estimate. The bottom plot (b) shows the distributions for the expected November 2021 recruitment estimate predicted from the regression given the 2021 survey result of 2.14. The 90% CIs are also shown.

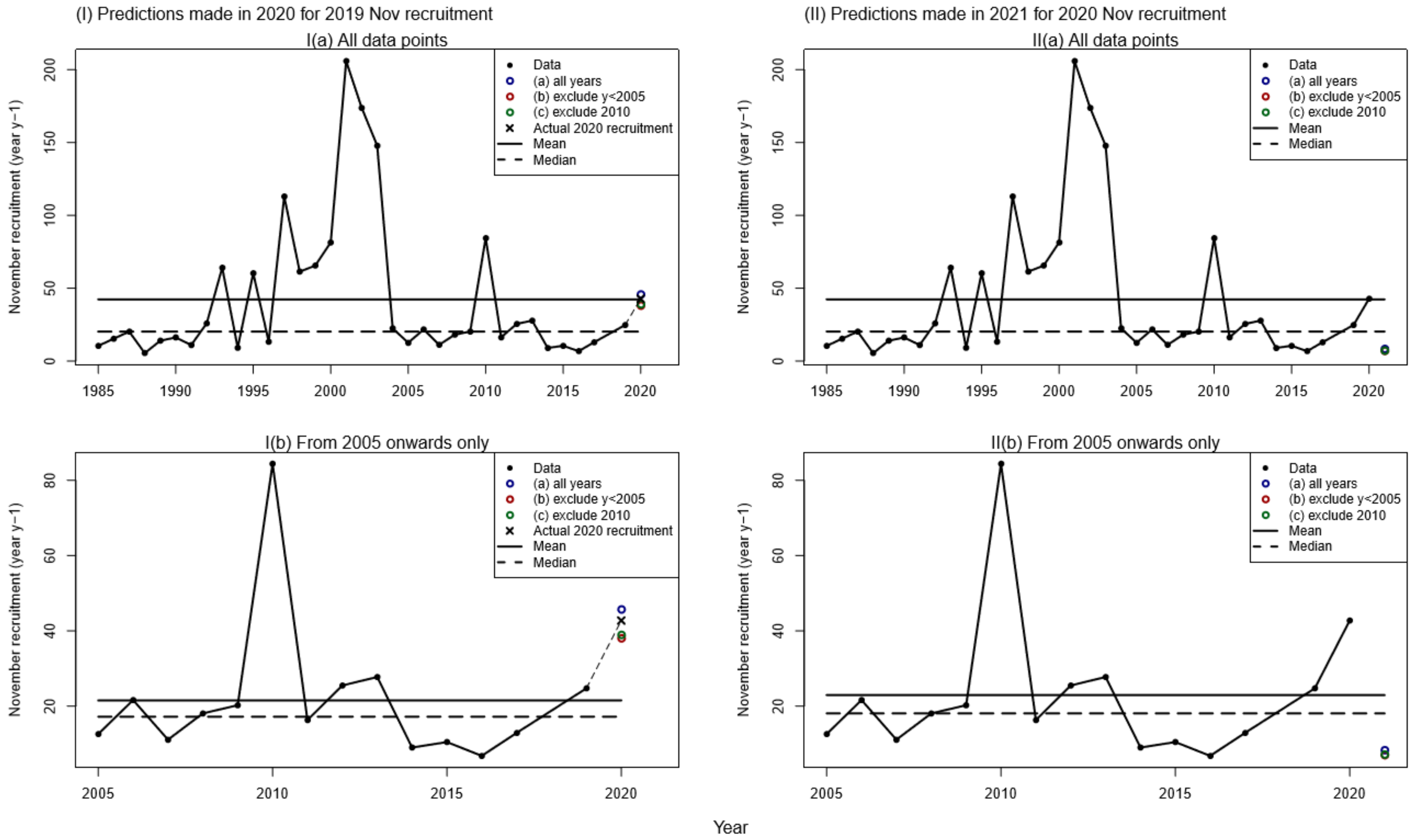

Figure 5: Plots of the November recruitment estimates against year, showing their mean and medians, for (I) the predictions that would have been made in 2020 based on the current recruitment series truncated at 2019 and (II) the current predictions for the 2020 November recruitment estimate based on the 2021 survey estimate. Results are shown for all data, and for 2005 onwards, to 2020. The means and medians on the left hand side plots (I) are to 2019, while for the right hand side plots (II) to 2020. Note that the November recruitment corresponds to the year prior to the corresponding value shown on the horizontal axis.