

SC/66a/SH/4

Simulation exercise to ascertain the
relative utility of collection of various data
types for informing future Southern
Hemisphere humpback whale
assessments and Breeding Stocks D, E1
and the Oceania stocks

Andrea Ross-Gillespie and Doug S. Butterworth



INTERNATIONAL
WHALING COMMISSION

Simulation exercise to ascertain the relative utility of collection of various data types for informing future Southern Hemisphere humpback whale assessments and Breeding Stocks D, E1 and the Oceania stocks

A. ROSS-GILLESPIE AND D.S. BUTTERWORTH¹

Contact e-mail: mlland028@uct.ac.za

ABSTRACT

The existing three-stock model for Breeding Stock D (West Australia or BSD), Breeding Stock E1 (East Australia or BSE1) and the collection of Oceania breeding sub-stocks (referred to as Breeding Stock O or BSO for convenience) is used to simulate future data which might be collected for these stocks, to ascertain which have the best potential to improve estimates of precision of key quantities associated with the population dynamics. For BSD, a future estimate of absolute abundance in 2017 and a series of relative abundance estimates each year from 2016 to 2020, where all estimates have CVs of 0.25, are considered. For BSE1, a similar further relative abundance series is considered, as well as new mark-recapture data collected over 2016 to 2020 where sighting probabilities are taken to be half the average achieved previously. Sighting probabilities are dealt with in a similar manner for future mark-recapture data considered for BSO. Both the further absolute abundance estimate for BSD and especially more mark-recapture data for BSO show some potential for providing improved precision in parameter estimates. However, there seems to be little increase in precision to be gained from further relative abundance data, bearing in mind that these results presume the current three-stock model to be correct. The three-stock model is further used to estimate the range of future observations that would remain compatible with that model's assumptions. Actual data collected in future could be compared to such ranges to check on the model's ability to continue to reflect reality.

INTRODUCTION

In 2014, the comprehensive assessment of the Southern Hemisphere humpback whale (initiated by the International Whaling Commission (IWC) over ten years previously) was completed. Currently, a synthesis of the outcomes of the comprehensive assessment is being developed. Looking to the future, it would be useful to know what sort of data would provide the most information to improve the results from the assessment models, in order to provide guidance of where resources should be invested in further research. The three main data sources used in humpback assessments are absolute abundance estimates, relative abundance estimates (trend information) and mark-recapture data.

This paper reports on results of a simulation exercise performed based on the 2014 three-stock model for West Australia (Breeding Stock D or BSD), East Australia (Breeding Stock E1 or BSE1) and the collection of Oceania breeding sub-stocks (referred to as Breeding Stock Oceania or BSO for convenience). The aim is to use the existing model and input data and to simulate additional future data in order to assess by how much these simulated data lead to a decrease in the current uncertainty in the values of selected population parameters.

A second aspect of this paper looks at providing a range of possible future abundance estimates and trend information that could be expected to be observed in the future given the current model outputs. When future observations become available, they can be compared to these ranges. If observations fall outside the ranges, this would indicate that the current model does not represent reality adequately and would need to be re-evaluated.

METHODS

Note that all computations were performed in R.

The 2014 three-stock model (IWC 2014)

The 2014 three-stock model defines two core feeding grounds: a western feeding ground between 70°E and 130°E and an eastern feeding ground between 130°E and 110°W (Figure 1). BSD and a proportion (γ) of the BSE1 population feed in the western feeding ground, while BSO and the remaining proportion ($1-\gamma$) of the BSE1 population feed in the eastern feeding ground. The historical feeding ground catches from 70°E to 110°W are allocated to the three stocks in proportion to the model-predicted abundances in each region. Note that breeding ground catches from New Zealand are split in proportion to the model-predicted BSE1 and BSO abundances.

¹ MARAM (Marine Resource Assessment and Management Group), Department of Mathematics and Applied Mathematics, University of Cape Town, Rondebosch, 7701, South Africa

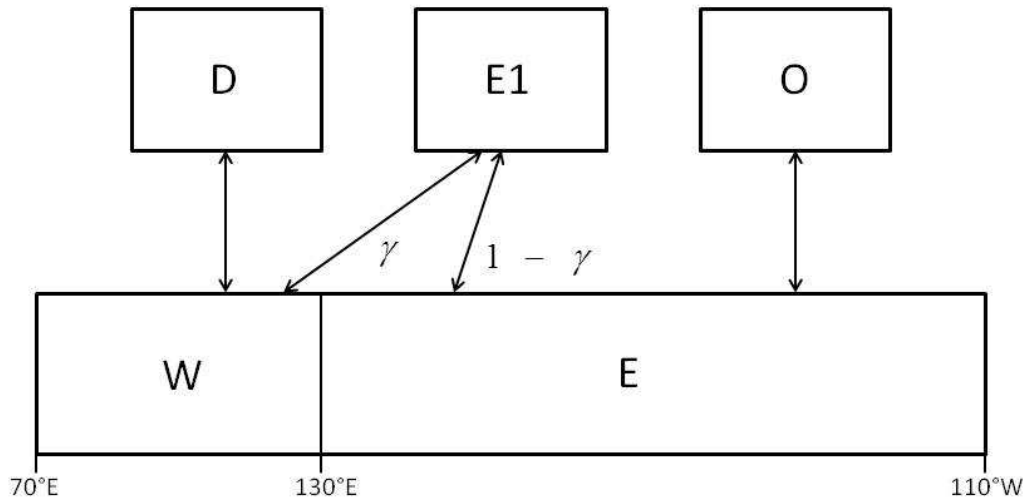


Figure 1: Diagrammatic representation of the 2014 three-stock model (IWC 2014).

The data utilised in the three-stock model are summarised in Table 1 below. Full details of the data are given in Ross-Gillespie *et al.* (2014).

Table 1: Data input for the 2014 three-stock model.

Stock	Data type	Source
BSD	Relative abundance	Hedley <i>et al.</i> (2011)
	Relative abundance	Bannister and Hedley (2001)
	N_{\min} constraint (3x no. of haplotypes)	3x27 (IWC 2014)
BSE1	Absolute abundance	Noad <i>et al.</i> (2011)
	Relative abundance	Noad <i>et al.</i> (2011)
	N_{\min} constraint (3x no. of haplotypes)	3x5 (IWC 2014)
BSO	Mark-recapture	Constantine <i>et al.</i> (2012)
	N_{\min} constraint (3x no. of haplotypes)	3x33 (IWC 2014)

Full details of the 2014 three-stock model assumptions can be found in section 3.1 of IWC (2014), and details of the population dynamics, Bayesian framework and the likelihood function can be found in Ross-Gillespie *et al.* (2014). Note that importance functions were utilised for BSD, BSE1 and BSO in order to improve sampling efficiency.

Simulation of ‘future’ data

Data set representing the 2014 three-stock model posterior distributions

In 2014, an initial set of 2 600 000 samples was generated following the sampling importance re-sampling or SIR algorithm (Rubin 1988). For each sample and each of the three populations, values of the population growth rate r and the log of a target abundance estimate $\ln N$ were drawn from prior distributions, and a value for the pristine carrying capacity K was found such that the model-predicted population trajectory matches the target abundance drawn, given the growth rate drawn and historical catches (see Ross-Gillespie *et al.* 2014 for more details). For the SIR algorithm, the normal procedure is then to use the likelihood function, prior distribution and importance functions for each sample to calculate the relative weight of each sample. In a re-sampling process 1000 equally-weighted samples are then chosen at random, generating what is considered to be a sample from the joint-posterior distribution. These 1000 values are used to calculate the median values and 90% probability intervals of population parameters of interest.

For the simulation exercise here, a slightly different approach has been taken. The likelihood function was evaluated for the 2 600 000 samples (from 2014) for the input data given in Table 1. Samples were sorted according to their likelihood weights, these weights were scaled so that their sum was one, and the cumulative weights computed. Samples whose cumulative weight contributed less than 0.001 (0.1%) to the total weight were discarded. This left

14847 samples², which were considered an adequate representation of the posterior distribution given the original data. Note that the samples from this reduced set are unequally weighted (in contrast to the equally weighted set of 1000 values mentioned above), and medians, 5th and 95th percentiles have to be derived in a slightly different manner, as described below.

The main motivation behind this approach was to be able to utilise the same base set of samples to provide comparable results when adding different simulated ‘future’ data sets. A secondary motivation was to improve the efficiency of the computer program as the time-consuming re-sampling process was removed.

Absolute abundance estimate

‘Future’ absolute abundance estimates were simulated for BSD for the year 2017. 100 samples were drawn at random from the reduced set of 14847, with the probability of choosing any particular sample being proportional to the likelihood weight (in terms of the original data) of that sample. The model-predicted 2017 population estimates $N_{2017}^{D,sample,i}$, obtained from the trajectories associated with the 100 samples, were used to generate 100 simulated absolute abundance estimates, $N_{2017}^{D,sim,i}$, by assuming that the log of each simulated abundance estimate is normally distributed about the log of the corresponding 2017 population estimate with a SE^3 of 0.25 to approximate a CV of the estimate itself of 25%, i.e. $\ln N_{2017}^{D,sim,i} \sim N(\ln N_{2017}^{D,sample,i}, 0.25^2)$. This CV value was chosen as likely typical of what might be achievable in practice. The likelihood contribution, evaluated for each of the 100 simulated abundance estimates and each of the 14847 samples, is given by

$$-\ln L_{i,j} = \frac{1}{2(0.25)^2} (\ln N_{2017}^{D,sim,i} - \ln N_{2017}^{D,sample,j})^2 \quad \text{for } i \in [1,100], j \in [1,14587] \quad (1)$$

Relative abundance estimates

Relative abundance series were simulated for BSD and BSE1, respectively. Similar to the absolute abundance estimate, the relative abundance estimates were generated by drawing 100 samples from the reduced 14847 set, taking the 2016-2020 population abundance values from the corresponding trajectories and adding noise by assuming that the simulated estimates are log-normally distributed about the values from trajectories each with a CV of 0.25 (again chosen as likely typical of what might be achievable in practice). The likelihood contribution for each of the 100 simulated series and each of the 14847 samples is given by

$$-\ln L_{i,j} = \bar{n} \ln \sigma_{i,j} + \frac{1}{2\sigma_{i,j}^2} \sum_y (\ln N_y^{sim,i} - \ln q_{i,j} - \ln N_y^{sample,j})^2 \quad \text{for } i \in [1,100], j \in [1,14587] \quad (2)$$

where

\bar{n} is the number of data points in the simulated series (i.e. five, from 2016-2020), and
 $q_{i,j}$ is a constant of proportionality, estimated by its maximum likelihood value:

$$\ln \hat{q}_{i,j} = 1/\bar{n} \sum_y (\ln N_y^{sim,i} - \ln N_y^{sample,j}) \quad \text{for } i \in [1,100], j \in [1,14587] \quad (3)$$

The σ parameter is the residual standard deviation, which is estimated in the fitting procedure by its maximum likelihood value:

$$\hat{\sigma}_{i,j} = \sqrt{1/\bar{n} \sum_y (\ln N_y^{sim,i} - \ln q_{i,j} - \ln N_y^{sample,j})^2} \quad \text{for } i \in [1,100], j \in [1,14587] \quad (4)$$

² The fact that only 14587 out of 2 600 000 samples constitute 99.9% of the total weight is largely a result of the highly informative BSE1 trend information, which strongly supports a high value of the BSE1 population growth parameter r . This means that any value of r^{EI} smaller than 0.103 has a very small chance of being re-sampled, leading to great sampling inefficiency. An importance function was used to help address the sampling-inefficiency, but did not eliminate the problem completely. A larger sample than 2 600 000 would be desirable, but was prevented by computing limitations.

³ If N is assumed to be log-normally distributed, then $\ln N$ is normally distributed with some mean μ and standard deviation σ . The median value of N is then e^μ while the CV of N is given by $\sqrt{e^{\sigma^2} - 1}$. Since the CV of N is relatively small, σ has been approximated here by the value of the CV of N .

Mark-recapture data

Mark-recapture data were simulated for BSE1 and BSO. Again, 100 samples were drawn from the 14847 and relevant model outputs extracted.

Mark-recapture data give:

n_y , the number of animals captured in year y , and

$m_{y,y'}$, the number of animals captured in year y that were recaptured in year y' .

These were the two quantities that needed to be simulated. However n_y is needed only to compute p_y , the probability that an animal is seen in a region in year y , given by:

$$n_y = p_y N_y \quad (5)$$

Rather than trying to simulate n_y , the p_y values were taken directly from model outputs: for each of the 14847 samples, $p_y^{sample,j}$ was taken to be the half of the average of model-predicted p_y values for the last five years for which mark-recapture data are available⁴. The p value was divided by two since the population size is likely to have more than doubled since the time of the existing data sets, from which the model-predicted p values were derived. Assuming that the same effort is put into sampling, the probability of re-sighting a whale will now be smaller. The factor is two is somewhat arbitrary, but serves to illustrate the impact of adding further mark-recapture data.

Note that for BSE1, the p values estimated from the Forestell *et al.* (2011) mark-recapture data were used, since the model fits this data set best. For BSO, the p values were taken from model estimates from the Jackson *et al.* (2012) males only mark-recapture data set. Since the males-only data set was used for BSO, Equation 5 above was divided by a factor of two for the BSO simulation, assuming gender parity.

The model predicted number of animals captured in year y that were recaptured in year y' , $\hat{m}_{y,y'}^{sample,j}$, is given by:

$$\hat{m}_{y,y'}^{sample,j} = p_y^{sample,j} p_{y'}^{sample,j} N_y^{sample,j} e^{-M(y'-y)} \quad (6)$$

M is the natural mortality rate (set here to equal 0.03 yr^{-1} as recommended by the IWC SC). $\hat{m}_{y,y'}^{sample,i}$ can thus be calculated for the 100 samples drawn for simulation purposes.

In order to simulate $m_{y,y'}$ values, these were assumed to be Poisson distributed with parameters $\hat{m}_{y,y'}^{sample,i}$. For any given sample of the 14847 and for each $[y, y']$ combination, $m_{y,y'}^{sim,i}$ was set to the values $k=1, 2, 3, \dots, 30$, and the corresponding Poisson probabilities were calculated:

$$f(k; \hat{m}_{y,y'}^{sample,j}) = P(m_{y,y'}^{sim,i} = k) = \frac{(\hat{m}_{y,y'}^{sample,j})^k}{k!} e^{-\hat{m}_{y,y'}^{sample,j}} \quad (7)$$

A value for $m_{y,y'}^{sim,i}$ was then chosen at random, with the probability of choosing each value of $k=1, 2, \dots, 30$ being proportional to the above probabilities.

The likelihood contribution, evaluated for each of the 100 simulated abundance estimates and each of the 14847 samples, is given by

⁴ Note that taking the average implies that $p_y^{sample,j} = p_{y'}^{sample,j}$ for all $[y, y']$ combinations. The approach for including mark-recapture data in the likelihood considers only pairs of re-sightings, i.e. each re-sighting is paired with the last sighting, but then considered a new sighting to be paired with the next sighting. It essentially assumes that whale has the same probability of being re-sighted no matter how many times it has been seen before. In light of this, setting $p_y^{sample,j} = p_{y'}^{sample,j}$ seems reasonable.

$$-\ln L_{i,j} = \sum_{y=y_0}^{y_f-1} \sum_{y'=y+1}^{y_f} [-m_{y,y'}^{sim,i} \ln \hat{m}_{y,y'}^{sample,j} + \hat{m}_{y,y'}^{sample,j}] \quad \text{for } i \in [1,100], j \in [1,14587] \quad (8)$$

Note that this simulation approach assumes that whales seen in earlier existing mark-recapture studies will not be seen again in the simulated data. In other words, for simplicity the current approach has not included consideration of existing sets of mark-recapture data, though resightings of those whales would be less likely in the future as a considerable period has elapsed since they were first seen.

Note that for BSE1, the original model does not fit directly to mark-recapture data. In order to compare how the mark-recapture improves the precision, the original model was re-run fitting to the Forestell *et al.* (2011) mark-recapture data instead of the Noad *et al.* (2011) relative abundance estimate. Thus for the BSE1 mark-recapture simulation, the 2 600 000 sample set was reduced according to this new likelihood. This resulted in a reduced sample set of 2256 (rather than 14847 as before).

Incorporating simulated data

For any given data type (absolute abundance/ relative abundance/ mark-recapture), the likelihood component of the new data was evaluated for each of the 14847 samples and the 100 simulated data sets, generating a bigger set of 1 458 700 unequally weighted samples (or 225 600 when using the Forestell *et al.* (2011) data, as per above). In order to compute medians, 5th and 95th percentiles for each parameter of interest (r , K , N_{2015}/K), the 1 458 700 (or 225 600) values for that parameter were sorted, and the corresponding likelihoods adjusted to sum to one, and the cumulative likelihoods computed. The median value was then the sample at 50% of the total cumulative weight, the 5th percentile at 5% and the 95th at 95%.

The process of generating 100 data sets and computing medians was repeated 50 (for the absolute abundance estimate and mark-recapture data) or 100 (for the relative abundance data) times in order to obtain an estimate in the variability of the simulation exercise results. The means and standard errors of the means of these 50 or 100 repetitions are reported.

Determining ranges for possible future observations

As further data are collected in the future, a question of interest will be whether the observations fall within the range expected given the outputs of the current three-stock model. If they do not, then this would be an indication that the three-stock model is mis-specified and does not adequately capture reality. The task is thus to obtain a range of values compatible with the current model with which future observations can be compared. This was achieved through a two step process. First a range of “true” values was computed by taking the model outputs of interest from the 14847 reduced sample set and computing medians, 5th and 95th percentiles. This is the range one would expect if the model is true, and there is no survey sampling error. Secondly a range of “expected” values for possible observations taking sampling error into account (assuming a CV of 0.25 as survey error) was generated. For this, the median, 5th and 95th percentiles were calculated for the 100 simulated data sets. The whole process was repeated 50 times and the averages of the 50 median, 5th and 95th percentiles values were calculated.

The ranges were calculated for the 2017 (absolute) abundance estimate (for BSD), the trend derived from the log of the 2016-2020 relative abundance estimates (for BSD and BSE1), as well as mark-recapture data for 2016-2020 (BSE1 and BSO). Ranges for absolute and trend estimates were not derived for BSO, since in the past this stock has been primarily monitored by use of mark-recapture data.

For the relative abundance estimates, the log-linear trend was calculated rather than working with the individual estimates since different relative abundance series are not directly comparable in absolute terms, but can be compared in terms of the trend obtained by fitting a regression through the log of the estimates.

For the mark-recapture data, ranges for $m_{y,y'}$, the number of animals that could be expected to be resighted in year y' after first being captured in year y , were calculated. “True” ranges (assuming no sampling error) were derived from Equation 6, while expected ranges (assuming sampling error) were derived from Equation 7. Results are given in terms of $\sum_y m_{y,y'}$, i.e. the total number of whales expected to be resighted in year y' (regardless of when they were first sighted), as well as $n_{y'}$, the total number of whales expected to be sighted in year y' (from Equation 5). It is important

to note that in order to provide these ranges, an assumption needs to be made that the sighting probabilities p_y (taken to be half of the model estimates for previous data sets) stay the same as used to compute the ranges.

RESULTS AND DISCUSSION

Data simulation exercise

Table 2(a)-(c) give the results for the data simulation exercise. Absolute abundance estimates were simulated for BSD, relative abundance series for BSD and BSE1 and mark-recapture data for BSE1 and BSO. For BSE1, results are given for both the case where the model fits to the Noad *et al.* (2011) relative abundance series and where the model fits to the Forestell *et al.* (2011) mark-recapture data instead. The median values were not expected to change much (and in fact did not) between the different simulated data sets, since the data were simulated from the three-stock model output under the assumption that the model is correct. The pertinent question is to what extent, if any, do the additional data help reduce the uncertainty about values of the estimates for the key quantities r , K and N_{2015}/K . The column of interest in Table 2 is thus the “Range” column, and the extent to which these ranges decrease as more data are added.

For BSD, there was a noticeable decrease in the range when simulated absolute abundance data were added, but interestingly not much difference when simulated relative abundance series were taken into account. This could be due to the fact that the model estimates the BSD population to be close to its pristine level (>95% thereof) in the years for which the data were simulated (2016-2020). The population trajectory is thus flattening out, and the simulated relative abundance series will show virtually as many negative as positive trends, thus not in fact adding much information to the existing data set in regard to the values of these quantities.

Similarly for BSE1, not much improvement was noticeable when further relative abundance estimates were added. The Noad *et al.* (2011) relative abundance series could be having an impact on this result, as it is highly informative and the associated parameter estimates already have fairly small uncertainty given the original data set, not leaving much room for improvement.

For both BSE1 and BSO, simulated mark-recapture data served to reduce the ranges of the key parameter estimates appreciably. The p values (i.e. the probability of sighting a whale in any given year for the five future years) obtained from model estimates for earlier years were half of 0.049 (90% probability interval [0.045, 0.052]) for BSE1 and half of 0.066 (90% probability interval [0.054, 0.074]) for BSO.

In summary, both an absolute abundance estimate for BSD and especially mark-recapture data for BSE1 and BSO show some potential for providing improved precision in estimates of key quantities. However, there seems to be little increase in precision to be gained from further relative abundance data, although other factors (such as the underlying assumption that the model is correct and the model-estimated level of recovery of the BSD population) would need to be considered when making decisions about what data sources would be preferred for future data collection.

Range for possible future observations

Figure 2 illustrates the ranges (5th to 95th percentile) of two observable quantities that could be expected to be estimated in future. In all cases a “True” range is given, which is the range we would expect if the model is correct and there is no survey sampling error. The “expected” range is the range that can be expected when sampling error (CV of 0.25) is taken into account and as such represents the range that would be pertinent for comparisons with possible future data, to determine whether the underlying model remains acceptable. This exercise serves to demonstrate the ability to generate expected ranges given the model assumptions, so that future observations can be compared to these plots in order to check if they fall within a range compatible with the existing model. If they do not, then this would be an indication that the three-stock model is mis-specified. It should be noted that when data become available in the future, these ranges would need to be re-evaluated given the actual sampling CVs and mark-recapture sighting probabilities, before valid comparisons can be made.

ACKNOWLEDGEMENTS

Funding support for this work from the International Whaling Commission is gratefully acknowledged. Comments from J. Jackson on an initial draft of this paper were appreciated.

REFERENCES

- Bannister, J.L. and Hedley, S.L. 2001. Southern Hemisphere Group IV humpback whales: their status from recent aerial survey. *Memoirs of the Queensland Museum* 47, 587-598.
- Constantine R, Jackson JA, Steel D, *et al.* 2012. Abundance of humpback whales in Oceania using photo-identification and microsatellite genotyping. *Marine Ecology Progress Series* 453, 249-261.
- Forestell, P.H., Kaufman, G.D. and Chaloupka, M. 2011. Long term trends in abundance of humpback whales in Hervey Bay, Australia. *Journal of Cetecean Research and Management (Special Issue)* 3, 237-241.
- Hedley, S.L., Dunlop, R.A. and Bannister, J.L. 2011. Evaluation of WA Humpback surveys 1999, 2005, 2008: Where to from here? In: Report to the Australian Marine Mammal Centre on work done to 6th May, 2011, p. 28.
- International Whaling Commission 2014. Report of the Scientific Committee, Annex H: Report of the Sub-Committee on other Southern Hemisphere whale stocks. IWC document IWC/65/Rep01 (2014). Bled, Slovenia, 12-24 May 2014 (<https://iwc.int/scientific-committee-reports>).
- Jackson, J.A., Anderson, M., Steel, D.S., Brooks, L., Braverstock, P., Burns, D., Clapham, P., Constantine, R., Franklin, W., Franklin, T., Garrigue, C., Hauser, N., Paton, D., Poole, M., Baker, C.S. 2012. Multistate measurements of genotype interchange between East Australia and Oceania (IWC breeding sub-stocks E1, E2, E3 and F2) between 1999 and 2004. Paper SC/64/SH22 presented to the IWC Scientific Committee, June 2012 (unpublished). 16pp
- Noad, M., Dunlop, R.A. and Paton, D. 2011. Abundance estimates of the east Australian humpback whale population: 2010 survey and update. In: Paper SC/63/SH22 presented to the IWC Scientific Committee, May 2011 (unpublished). 12pp.
- R Core Team. 2013. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org/>.
- Ross-Gillespie, A., Butterworth, D.S. and Holloway, S.J. 2014. Assessment results for humpback breeding stocks D, E1 and Oceania following recommendations from SC 65a. IWC document SC/65b/SH04rev.
- Rubin, D.B. 1988. Using the SIR algorithm to simulate posterior distributions. P. 395-402 in Bernardo, J.M., DeGroot, M.H., Lindley, D.V. and Smith, A.F.M. (ed.). 1988. Bayesian Statistics 3: Proceedings of the third Valencia International Meeting, June 1-5,1987. Clarendon Press, Oxford. 805pp.

Table 2 : Results of the simulation exercise where the “Original data set” refers to the reduced sample set of 14847⁵ with the original historical data as given in Table 1. The other rows contain the results for when additional simulated data were incorporated and a set of 14847x100 likelihood values with associated parameter values was constructed. Both the above-mentioned sample sets are unequally weighted. Medians and percentiles were computed by sorting the parameter values for the parameter in question, adjusting the corresponding likelihood values to sum to one and computing the cumulative likelihoods. The median value is then the sample at 50% of the total cumulative weight, the 5th percentile at 5% and the 95th at 95%. The simulation process was repeated 50 times for the absolute abundance estimate and mark-recapture data, and 100 times for the relative abundance data, and the tables below give the means of the medians and percentiles over the 50 or 100 repetitions. Standard errors of the mean are given in parenthesis. The last column gives the range (5th-95th percentile) of the parameter estimates. Note that for BSE1, results are also given for when the model is fit to the Forestell *et al.* (2011) mark-recapture data, instead of the Noad *et al.* (2011) relative abundance series.

(a) BSD

(i) r^D	Median	5th percentile	95th percentile	Range
Original data set	0.0904 (0)	0.0524 (0)	0.104 (0)	0.0513 (0)
Additional absolute abundance	0.0909 (0.00006)	0.0544 (0.00538)	0.104 (0.00026)	0.0493 (0.00004)
Additional relative abundance	0.0899 (0.00038)	0.0519 (0.00064)	0.104 (0.00005)	0.0517 (0.00060)
(ii) K^D	Median	5th percentile	95th percentile	Range
Original data set	21669 (0)	19016 (0)	29865 (0)	10849 (0)
Additional absolute abundance	21617 (8)	19028 (2)	28661 (21)	9633 (19)
Additional relative abundance	21696 (53)	18992 (18)	29892 (193)	10900 (178)
(iii) N_{2015}/K	Median	5th percentile	95th percentile	Range
Original data set	0.948 (0)	0.788 (0)	0.991 (0)	0.204 (0)
Additional absolute abundance	0.948 (0.00000)	0.794 (0.00295)	0.991 (0.00000)	0.197 (0.00033)
Additional relative abundance	0.948 (0.00125)	0.791 (0.06727)	0.991 (0.00036)	0.200 (0.00725)

(b) BSE1

(a) r^{E1}	Median	5th percentile	95th percentile	Range
Original data set	0.1053 (0)	0.1030 (0)	0.1060 (0)	0.0029 (0)
Additional relative abundance	0.1053 (0.00000)	0.1030 (0.00001)	0.1060 (0.00000)	0.0029 (0.00001)
Additional mark-recapture	0.1053 (0.00000)	0.1034 (0.00000)	0.1059 (0.00000)	0.0025 (0.00000)
Original (fit to Forestell data)	0.1056 (0)	0.1042 (0)	0.1059 (0)	0.0016 (0)
Additional mark-recapture	0.1055 (0.00000)	0.1042 (0.00000)	0.1059 (0.00000)	0.0017 (0.00000)
(b) K^{E1}	Median	5th percentile	95th percentile	Range
Original data set	26253 (0)	21864 (0)	29153 (0)	7289 (0)
Additional relative abundance	26171 (51)	21805 (70)	29086 (25)	7281 (56)
Additional mark-recapture	26648 (15)	23688 (35)	28909 (12)	5221 (45)
Original (fit to Forestell data)	26522 (0)	22831 (0)	28402 (0)	5572 (0)
Additional mark-recapture	26621 (22)	24490 (43)	28735 (8)	4245 (39)
(c) N_{2015}/K	Median	5th percentile	95th percentile	Range
Original data set	0.759 (0)	0.685 (0)	0.841 (0)	0.155 (0)
Additional relative abundance	0.761 (0.00162)	0.690 (0.00127)	0.842 (0.00150)	0.152 (0.00117)
Additional mark-recapture	0.702 (0.00137)	0.656 (0.00128)	0.753 (0.00151)	0.098 (0.00047)
Original (fit to Forestell data)	0.706 (0)	0.671 (0)	0.773 (0)	0.103 (0)
Additional mark-recapture	0.702 (0.00005)	0.656 (0.00003)	0.732 (0.00011)	0.076 (0.00008)

(c) BSO

(i) r^O	Median	5th percentile	95th percentile	Range
Original data set	0.0914 (0)	0.0721 (0)	0.101 (0)	0.0293 (0)
Additional mark-recapture	0.0935 (0.00009)	0.0809 (0.00015)	0.101 (0.00002)	0.0202 (0.00016)
(ii) K^O	Median	5th percentile	95th percentile	Range
Original data set	14641 (0)	10106 (0)	20105 (0)	9999 (0)
Additional mark-recapture	13854 (18)	10233 (13)	18770 (13)	8537 (22)
(iii) N_{2015}/K	Median	5th percentile	95th percentile	Range
Original data set	0.440 (0)	0.283 (0)	0.656 (0)	0.373 (0)
Additional mark-recapture	0.474 (0.00041)	0.335 (0.00065)	0.639 (0.00136)	0.304 (0.00195)

⁵ 2256 samples were used for the case where BSE1 was fit to the Forestell *et al.* (2011) mark-recapture data instead of the Noad *et al.* (2011) relative abundance series.

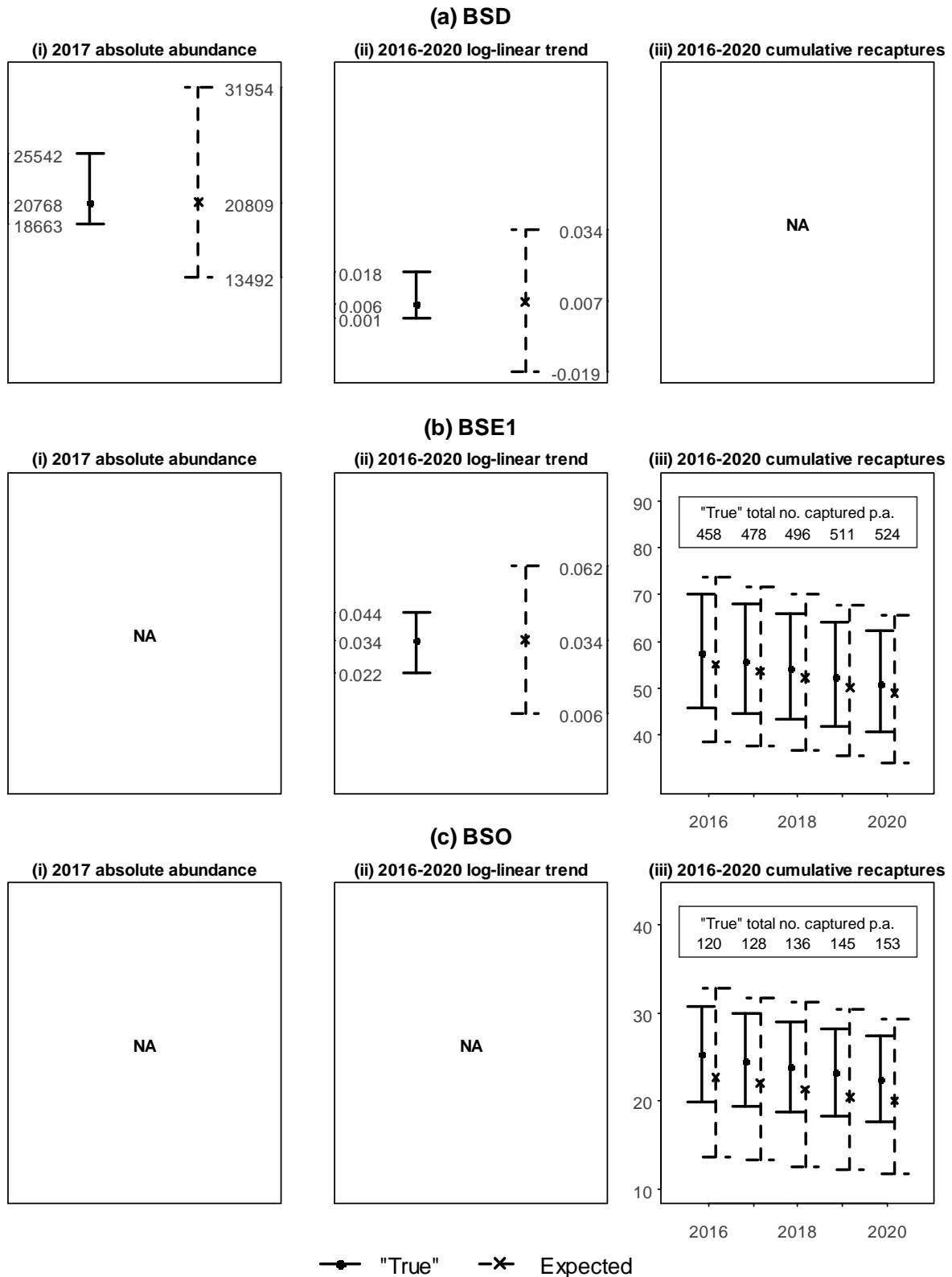


Figure 2: Ranges of possible future observations under the assumption that the current three-stock model is correct. “True” corresponds the median and percentiles derived from the 14847 samples developed from the original data set, while the medians and percentiles for the “Expected” ranges are derived from the simulated data sets, which take future survey sampling error into account, averaged over 50 replicate simulations. Ranges for absolute abundance estimates and log-linear trend are given for BSD; ranges for log-linear trend and cumulative recaptures for BSE1, and ranges for cumulative recaptures for BSO. ‘NA’ indicates that the corresponding data source has not been associated with the stock in question historically. The mark-recapture plots show the “true” and expected number of recaptures that can be expected each year from 2016-2020, assuming that the sighting probabilities are half of what they were for the last five years of the mark-recapture data sets (half of $p=0.049$ [0.045, 0.052] for BSE1 and half of $p=0.066$ [0.054, 0.074] for BSO). The median total number of captures to be expected each year (Equation 5) is also given. Note that the ranges for BSE1 cumulative recaptures are for the case where the model is fit to the Noad *et al.* (2011) relative abundance series.