CCSBT-ESC/1809/14
(ESC Agenda item 8.1)

OCEANS AND ATMOSPHERE
www.csiro.au

Close-kin project report

Campbell Davies, Mark Bravington, Peter Grewe and Paige Eveson
CCSBT-ESC/1809/14

Citation

Davies, C.R., Bravington, M.V., Grewe, P.D. Eveson, P.J., Lansdell, M. Hill, P. and Aulich, J. (2018) Close-kin project report. CCSBT-ESC/1809/14. CSIRO, Australia.

Copyright

Important disclaimer

Contents

Acknowledgements

Thanks are due to many who made this work possible.

The Research Institute for Tuna Fisheries (Bali) for collaboration on sampling adults from spawning ground.

The late Tony Jones (Protec Marine P/L) and Adam Kemp at Seatec P/L, for sampling of juveniles; and the Australian Southern Blue□n Tuna Industry Association (ASBTIA) for ongoing support.

CCSBT and FRDC for funding.

Andrzej Kilian and all at Dart Pty Ltd for DArTcap genotyping;

Matt Landsdell, Jess Farley, Peta Hill, Floriaan Devloo-Delva and Jorden Aulich for for great e□orts in logistics to complete all the processing in time.

Abstract

This paper reports on the completion of a second stand-alone Close-Kin Mark Recature model for SBT that uses Parent Offspring (POP) and Half-sibling Pairs (HSP) identified using specifically designed SNiP assays for SBT and samples of adults and juveniles collected between 2006 and 2015. DNA was extracted from ~ 17,000 individuals with a total of ~15,000 individuals (4,238 adults and 10,952 juveniles) analyses for POP and HSP following DNA and genotyping quality control. A totoal of 77 POPs (including the 45 found in the original CCKMR study) and 140 de□nite HSPs and 4 Full-Sibling Pairs were identified. The true number of HSPs is estimated to be about 10% greater, because of the stringent criteria required to exclude false-positives. Examination of mitochondrial DNA indicates that about 65 of the 140 HSPs shared a mother whereas 75 shared a father, consistent with an equal sex-ratio in adult SBT. The stand-alone CKMR model used in the original study was extended to include: i) HSP, ii) the extended time series, iii) to allow selectivity to vary and iv) to free selectivity from fecundity, as well as some other minor revisions for consistency with the CCSBT Operating Models. The estimates of abundance from the new POP+HSP model and data are fairly similar to the values from the previous POP-only study, with the new estimates of SSB ae about 10% higher on average— a degree of change which is consistent with that expected from sampling variability, given there were 45 POPs available in the original study, whereas there are 76 in total in the updated data series. The overall summary statistics of biomass and numerical abundance varied relatively little across the model options explored, but there were differences in the age-specific components (n16p, nPLUS, Rcts), whereby models with estimated, rather than fixed, selectivity predict more old and fewer young adult fish. All options explored, with one exception, show very strong incoming cohorts of 8yo from about 2012 onwards and, by 2014, those cohorts have started to make an impact on overall SSB and TRO, so substantial upward trends in TRO and SSB would be expected from 2015 onwards as these recent adults continue to grow. A key difference from original study is that the HSPs are now providing a direct signal on overall adult *z*, and this seems broadly consistent with the overall *z* that was inferred under the assumptions of the POP-only model. However, the new model does show some preference for a somewhat higher survival for young adults, and an overall dome-shaped selectivity. This difference would have some effect on turnover rates, and estimated incoming 8yo recruitments. What is not yet clear is how seriously to take that dome-shape. Since the treatment of selectivity for the *LSfreq* data is still not fully satisfactory, especially with respect to the observed sex ratios and, hence, this warrants further investigation. Lastly, the best practical fits with estimated selectivity are consistent with $\alpha HSP = 1$. Since there is no strong *a priori* reason to expect $\alpha HSP < 1$ and there is no currrent evidence that it is below 1, despite reasonable sample sizes (numbers of POPs and HSPs), it seems fair to assume $\alpha HSP = 1$ for the Reference set of the CCSBT OMs until there is any clear evidence to the contrary. Periodic updates of the stand-alone CKMR model could be used to regularly to review this assumption.

1 Introduction

The initial SBT application of Close-kin Mark Recapture (CKMR) used specifically designed microsatellite loci (Bravington et al, 2016) to identify 45 Parent-Offspring-Pairs (POPs) in about 14,000 samples of known spawning adults (Indonesia) and known-age juveniles (Great Australian Bight). These were embedded in a statistical mark-recapture framework, and combined into a stand-alone mini-assessment of the adult component of the population that used length and age composition data from Indonesian longline catches on the spawning ground, plus histological information on relative daily fecundity-at-size (Farley and Davis 1999; Farley et al 2014). This stand-alone assessment was able to estimate a time-series of absolute spawning stock biomass, effective annual fecundity-at-size and total mortality rate of the mature component of the population. Full details of the sampling design, genotyping, quality control, procedures for identifying POPs, estimation model, and independent review process are provided in (Bravington et al 2014). The approach and the final results were reviewed by the CCSBT Extended Scientific Commitee in 2012 and 2013 and accepted as: (i) a valid fishery-independent estimate of spawning stock abundance for SBT, and (ii) as valid input data (i.e. the 45 POP information and associated comparisons) for the CCSBT OMs (Hillary et al 2012, 2013, Anon 2013).

1.1.1 Beyond Parent-Offspring-Pairs and microsatellites

The potential of CKMR for directly estimating absolute abundance and other key demographic parameters for natural resource management, has led to substantial investments in the theory and practice subsequent to the first tranche of SBT-related work. This has included:

1. Development of demographic CKMR models that can use Half-Sibling Pairs (HSPs): where two animals have one shared parent, as well as Parent-Offspring Pairs (Bravington et al 2016a).
2. Reviewing and testing the suitability and cost-effectiveness of different Next Generation Sequencing platforms (e.g. DArT, RadSeq, Sequenom, GBS) for large-scale close-kin genotyping to find HSPs and POPs.
3. Development of general statistical/demographical theory for CKMR (Bravington et al 2016a)
4. Design and implementation of CKMR studies for other species (especially sharks) with very different sampling and demography (e.g. where only juveniles can be sampled) (e.g Hillary et al 2018)
5. Design work for CKMR as a long-term monitoring tool for CCSBT, using HSPs as well as POPs (Bravington and Davies 2013, Bravington 2014 and Bravington et al 2015). The long-term use of CKMR, along with gene-tagging to estimate recruitment (Preece et al 2015), is now endorsed by the CCSBT Scientific Committee and funded under the CCSBT Scientific Research Program.

In addition to the fundamental development work referred to above, the results reported here draw directly on investments and outputs from three related projects:

1. Long-term monitoring of Indonesian catches from the SBT spawning grounds – CSIRO-Agency for Marine and Fisheries Research, Indonesia

Landings of SBT from Indonesian vessels fishing on the spawning grounds south of Bali-Java have been monitored since 1993, with otoliths collected from 1994, as part of a collaborative program between the research agency of the Ministry for Marine Affairs and Fisheries of Indonesia (Farley et al 2014, Farley et al 2015, Farley et al 2017). The sampling program was extended in 2006 to include tissue samples, providing the adult samples for the original CKMR study, and was been maintained through support form CSIRO, MMAF-Indonesia and DAWR and, now, CCSBT (see below).

2. Collection and genotyping of 2015-2016 SBT samples for close-kin – CCSBT-CSIRO

Since 2014, this project has supported the annual collection of tissue samples of adult SBT from Benoa, Indonesia, and 3 year old juveniles from Port Lincoln, Australia, as well as DNA extraction, archiving and sequencing of the DNA (Farley et al 2018).

3. Estimating abundance, mortality and selectivity using Close-kin pairs - CSIRO.

This project developed the SNP markers, assays and genotyping pipelines required to accurately and reliably identify Half-Sibling and Parent-Offspring pairs from samples of SBT DNA and the modelling framework to estimate spawning abundance, total mortality and selectivity using these two types of close-kin data (Bravington et al 2017). This modelling framework and analysis pipelines have been used in this project, along with the POPs and HSP identified from the genotyping, to generate a time series of absolute adult abundance, mortality and selectivity for SBT from 2002 through to 2014.

2 Methods

2.1 From tissue to kin

2.1.1 Sample collection, tissue processing, DNA extraction and archiving

Sampling of adults takes place in Benoa, Indonesia, during processing of catches from the longline fishery on the spawning ground, with tissue and otoliths collected at the same time. Sampling of juveniles caught by the Australian the purse seine fishery takes place in Port Lincoln, Australia. Sampling is done in the processing factories during harvest from grow out pens, some 3-6 months after capture. Samples consist of a biopsy containing ~300mg of tissue, which are placed in 2.0 mL cryovials, frozen, and transported to the CSIRO Marine Laboratories in Hobart.

Tissues are held at −80◦C until sub-sampled in preparation for DNA extraction. For each fish selected for subsampling, a ~15mg slice of tissue is weighed and placed into an extraction chamber for tissue digestion. An Eppendorf EP motion robot completes the DNA extraction and produces two final 96-well plates: a sequencing plate, and a replica DNA archive plate. Each plate contains DNA from 92 individuals, as well as two blanks and two control tissue samples; the position of which allow unique identification of each plate for quality control (QC) purposes. The archive plates are stored frozen at −80◦C where they remain unless required for further testing.

2.1.2 DNA sequencing

The sequencing plates are sent for sequencing at Diversity Array Technologies Pty Ltd (DArT https://www.diversityarrays.com), Canberra, using a specific variant of Genotyping- By-Sequencing designed by CSIRO and DArT for close-kin purposes, known as DArTcap. This involves laboratory pre-processing of the plates; analysis using a high-throughput sequencer and assays for a specific set of ~1500 SNP loci; and bioinformatic analysis of the terabytes of resulting data, to produce specific data summaries for each fish at each SNP locus of interest.

2.1.3 Genotyping for kin identification

The final step prior to kin identification takes the data summaries provided by DArT and turns them into multi-locus genotypes for each individual fish i.e., for each fish and each locus, the pair of alleles inferred to be present. This genotype-calling entails some quite complicated algorithms developed at CSIRO specifically for DArTcap sequencing data, and also estimates the genotyping error-rates for each locus. The latter is essential for robust identification of kin, in particular HSP, and associated uncertainty.

2.1.4 Kin Identification

The final step prior to the CKMR modelling itself is kin-finding, which is based on the inferred genotypes and the error-rates from the multi-locus geneotyping. For kin identification CSIRO has developed generic algorithms (i.e. not specific to DArTcap) from basic statistical principles (For those interested in more detail, these are summarized in the Appendix and in section 5 of Bravington et al. (2016b)). Control of rates of false-positive and false-negative kin is crucial to this process, since ~100,000,000 comparisons might be needed to find only ~100 true kin-pairs.

2.1.5 Maternal and Paternal Half-sibling Pairs

it is possible to determine whether each HSP is the result of sharing the same father (Paternal HSP) or mother (Maternal HSP) using genetic analysis of mitochondrial DNA of samples of identified HSPs. The comparison of maternal and paternal HSP provides insights into differences in how fecundity varies with age between males and females, and on the true sex ratio of adults. The mtDNA data are incorporated directly in the standalone CKMR estimation model (see below).

Table 1: Summary of DNA extractions and samples successfully sequenced for SNPs by DArTcap for identification of Parent-Offspring-Pairs and Half-Sibling-Pairs funded by FRDC 2016-044.

Year/season	Source	Sampled	DNA Extracted	Sequenced
2006	Port Lincoln	4042	1472	1468
2007	Port Lincoln	4085	1472	1443
2008	Port Lincoln	4138	1564	1488
2009	Port Lincoln	4100	1473	1458
2010	Port Lincoln	4071	1472	1467
2011	Port Lincoln	4000	1012	1011
2012	Port Lincoln	4000	1012	1000
2013	Port Lincoln	1600	1012	998
2014	Port Lincoln	1600	1012	998
2005-06	Indonesia	216	0	0
2006-07	Indonesia	1520	0	0
2007-08	Indonesia	1594	0	0
2008-09	Indonesia	1637	0	0
2010-11	Indonesia	1013	1012	1011
2011-12	Indonesia	565	552	549
2012-13	Indonesia	1381	1012	998
2013-14	Indonesia	1642	1011	991
Total		**41204**	**15088**	**14880**

The CCSBT OMs are not sex structured and thus cannot use the maternal/paternal information in the HSPs; instead, it is assumed that there is little difference among adult male and females both in total numbers, and in how age and length affect individual fecundity. Thus, the standalone CKMR model, which does separate adults by sex and treats the MHSPs and PHSPs separately, can provide a test of the CCSBT OM assumptions.

Large-scale processing of archived tissue samples for DNA extraction began in October 2016, following in principle agreement between CSIRO and FRDC on funding arrangements. All DNA extractions were complete by January 2017 and sequencing of ~ 16,000 fish at DArT began in February 2017; the full set of sequencing- files were received by CSIRO at the end of March 2017. In parallel, CSIRO developed quality control (QC), genotype-calling and kin-finding algorithms suitable for the new type of genetic data. From April to June, these algorithms were refined and applied to deliver reliable sets of POPs and HSPs (Bravington 2017) suitable for use in the 2017 reconditioning of the CCSBT OMs (Hillary et al 2017b).

Table 2. Summary of DNA extractions and samples successfully sequenced for SNPs by DArTcap for identification of Parent-Offspring-Pairs and Half-Sibling-Pairs funded by related CCSBT project. * Note the original proposal was to include the 2014-15 Indonesian samples. However, these we not included due to tissue quality issues with this particular year. The 2009-10 samples were substituted in their place.

Year/season	Source	Sampled	DNA Extracted	Sequenced
2015	Port Lincoln	1600	1011	1005
2009-10	Indonesia	1200	1012	1012
2014-15*	Indonesia	1609	0	0
Total		**6009**	**2023**	**2017**

Table 3. Summary of final number of samples genotyped and used in kin-finding by location and year

Year/season	Indonesia (Adults)	Port Lincoln (Juveniles)
2006	0	1281
2007	0	1305
2008	0	1315
2009	0	1317
2010	943	1284
2011	931	938
2012	527	844
2013	933	873
2014	904	873
2015	0	922
Total	**4238**	**10952**

2.2 Incorporation of Parent-Offsping and Half-Sibling Pairs in CCSBT Operating Models for 2017 stock assessment

The Parent-Offspring-Pair data were available in time to be incorporated into the OM updates for June 2017 meeting of the Operating Model and Management Procedure (OMMP) Technical Group in preparation for the 2017 assessment of stock status (Hillary et al., 2017a). The HSPs were reported separately (Bravington, 2017). There was insufficient time to incorporate these data into the CCSBT OMs prior to OMMP8, due to the more complex nature of the analysis and quality control procedures, relative to the POPs. Notwithstanding this, the OMMP Technical Group, both data sets were available for the OMMP meeting to review, and the technical group recommended the new POP and HSP data be included in the 2017 assessment, conditional on review of the results of incorporation of the HSP data (Anon 2017a). The HSP data were incorporated into the OMs following the June meeting (Hillary et al. 2017b) and the results reviewed at a special webinar meeting of the OMMP Technical Group. The OMMP was satisfied with the results and agreed to include the HSP data series in the reference set of OMs for the 2017 stock assessment, and to an additional sensitivity run to examine the implications of a specific assumption relating to the HSP (Anon 2017b).

2.3 Overview of Stand-alone CKMR model with POPs and HSPs

There are four main changes from the stand-alone CKMR model used in the original application to SBT (Bravington et al 2014, 2016a). These are:

1. Extending the time-series of data through to 2014 (length⁄sex⁄age frequencies; genotypes);

2. Modifying the length-frequency model to allow for annual changes in selectivity;

3. Inclusion of Half-Sibling Pairs

4. “Freeing” selectivity from fecundity

There are also some minor differences associated with the different nature of the input data and providing greater consistency with assumptions of the CCSBT Operating Models, where it reasonable to do so (e.g. plus-group at age 30, rather than 25).

The most important change is the inclusion of HSPs. The key benefit of HSPs is that they lead rather directly to an estimate of average adult total mortality (z), or more accurately of the rate-of-turnover of Total Reproductive Output[1] . Parent-Offspring Pairs alone do not carry intrinsic

[1] (**T**otal **R**eproductive **O**utput, in some arbitrary but fixed unit such as “equivalent average 16yo SBT”— basically the same as SSB, except that TRO is a more accurate measure of what SSB is “trying” to measure.

information on z so. In the original application to SBT Bravington et al (2014) addressed this limitation by assuming :

- that selectivity was directly proportional to residence time on the spawning grounds;
- that male and female adult mortality rates were equal.

These assumptions could not be directly tested within the original framework, nor was there other information (e.g. electronic tagging data) with which to test it.

Incorporating HSPs into the log-likelihood of the stand-alone model is relatively straightforward. The underlying equations are unambiguous about how that should be done (see Appendix, and the explanation of HSP probabilities in section 3.9 ofBravington et al 2016b). The greater complication lies in allowing a more flexible selectivity-at-length relationship, plus allowing that to vary from year to year.

There are four main datasets included in the stand alone model, each of which contributes separately to the log-likelihood:

1. Parent-Offspring Pairs (POPs): The number of comparisons, and number of POPs found, broken down by year-of-adult-capture, adult sex, length and age, and year-of-juvenile birth. Note that comparisons are not made between adults caught in one season and juveniles born in that same season.
2. Half-Sibling Pairs (HSPs): The number of comparisons, and number of Maternal HSPs (shared mother) and Paternal PHSPs (shared father), broken down by birth-years of the two juveniles being compared[2]. Note, comparisons are not made between juveniles from the same cohort.
3. Age, given Length and Sex (A@LS) in Indonesian otolith subsamples: That is, given the length and sex of a fish, what age was the estimated age?
4. Length frequency composition of the monitored component of the catch from the spawning ground (LSfreq): Samples from Indonesian fishery that are selected for otolith extraction (though not all otoliths are subsequently read). These are assumed to constitute a random subsample of landings in Bali.

Note that the first three inputs are truly “fishery-independent”, in that they are driven only by fish biology and the qualitative circumstances of sampling, rather than by changes in fishing/fleet behaviour. There is a clear, albeit complicated, logical path for how to model POPs, HSPs, and A@LS statistically—- i.e., for how the demographic parameters set the statistical distributions of each. There is also no reason to expect these sub-models to change with time (except if there are cohort-specific growth changes).

However, the fourth dataset, LSfreq, is subject to selectivity of the Indonesian longline fishery; partly as a result of fish behaviour, and partly via boats fishing in different regions and⁄or with different gear setups (e.g. longline setting depth). This could be modelled in many different ways.

[2] It is possible to determine whether a HSP is by maternal or paternal decent by examination of the mitochondrial DNA, which is inherited from the mother only. See section 3.3 of appendix of Bravington et al (2017), for further detail.

The sample sizes are much bigger than for the other three datasets, so unless the intrinsic variability in selectivity can be successfully allowed for, this data will tend to dominate the likelihood and distort the fit.

The original SBT model (Bravington et al 2014) assumed that numbers-at-length-and-sex followed an over-dispersed multinomial distribution each year, with expected values predicted from the population dynamics and the estimated residence-time at length and sex. The extent of over-dispersion (corresponding to about an 8-fold effective reduction over actual sample size of LSfreq) was previously estimated based on empirical comparisons of length frequencies in adjacent years. The over-dispersion adjustment is applied to each centimetre length class separately; it is very simple, but fails to capture large-scale shifts in selectivity, whereby many adjacent length-classes may become over- or under-represented in a given year. Since 2010, such systematic shifts have been apparent in the length and age frequency data from the Indonesian spawning ground fishery(Farley et al 2017), and over-dispersion alone is clearly no longer adequate to account for this shift. Instead, a more complex annually-variable, spline-based selectivity adjustment, estimated via random effects, has been incorporated (see Appendix for details). A similar approach is used in the CCSBT OMs, albeit age-based rather than length-and-sex-based.

2.3.1 Exploring key model assumptions

The addition of HSP data allows a much wider range of model options to be explored than was possible with POPs alone. The CKMR model has been developed in a statistical framework (Restricted or Residual Maximum Likelihood (REML[3])), which is based on an explicit log-likelihood. This allows different "options" of the model, e.g., if survival rate is allowed to depend on sex as well as age, to be explored seamlessly, with accompanying diagnostics, to identify a preferred option from the set explored.

We have examined a reasonable number of options (see Appendix for full details) and the results indicate that the main issues to consider are:

- **selectivity:** constrained to match residence-time, as per Bravington et al (2014), or more flexible (e.g. allowing dome-shape)?
- **α HSP :** estimated (by sex), or fixed at 1? This parameter is there to allow for any unexpected discrepancy between the observed numbers of POPs and of HSPs, which could arise if some adults are systematically under-represented in the HSPs. For example, adults who systematically tend to breed offspring that migrate to South Africa in summer, rather than to the GAB. Ideally, we would like to see α HSP =1, but this needs to be checked because there may be some

[3] Restricted or Residual Maximum Likelihood— i.e., with random effects whose variance can be estimated consistently inside the model. The framework is basically Bayesian, although it is operationally not essential to use MCMC to estimate parameters, nor to put explicit priors on every parameter (and we have not done so here).

surprises in the reproductive biology. It is logically impossible to have α HSP >1 (see Appendix for details), although an *estimate* might come out slightly higher than 1 just by chance.

- **LSfreq weighting:** should the LSfreq be used at full strength, or down-weighted? If the selectivity model is adequate, then there should be no need to apply an overall down-weighting to the LSfreq; but if it is not, then an option with down-weighted LSfreq may be more robust.

Almost all options described here fit the CKMR data and the A@LS data quite well, but there remains some tensions with the LSfreq data, which warrant further investigation; in particular, no option can yet accurately reproduce the changing in observed sex frequency.

3 Results and Discussion

3.1 Kin Identification

3.1.1 Parent Off-spring Pairs

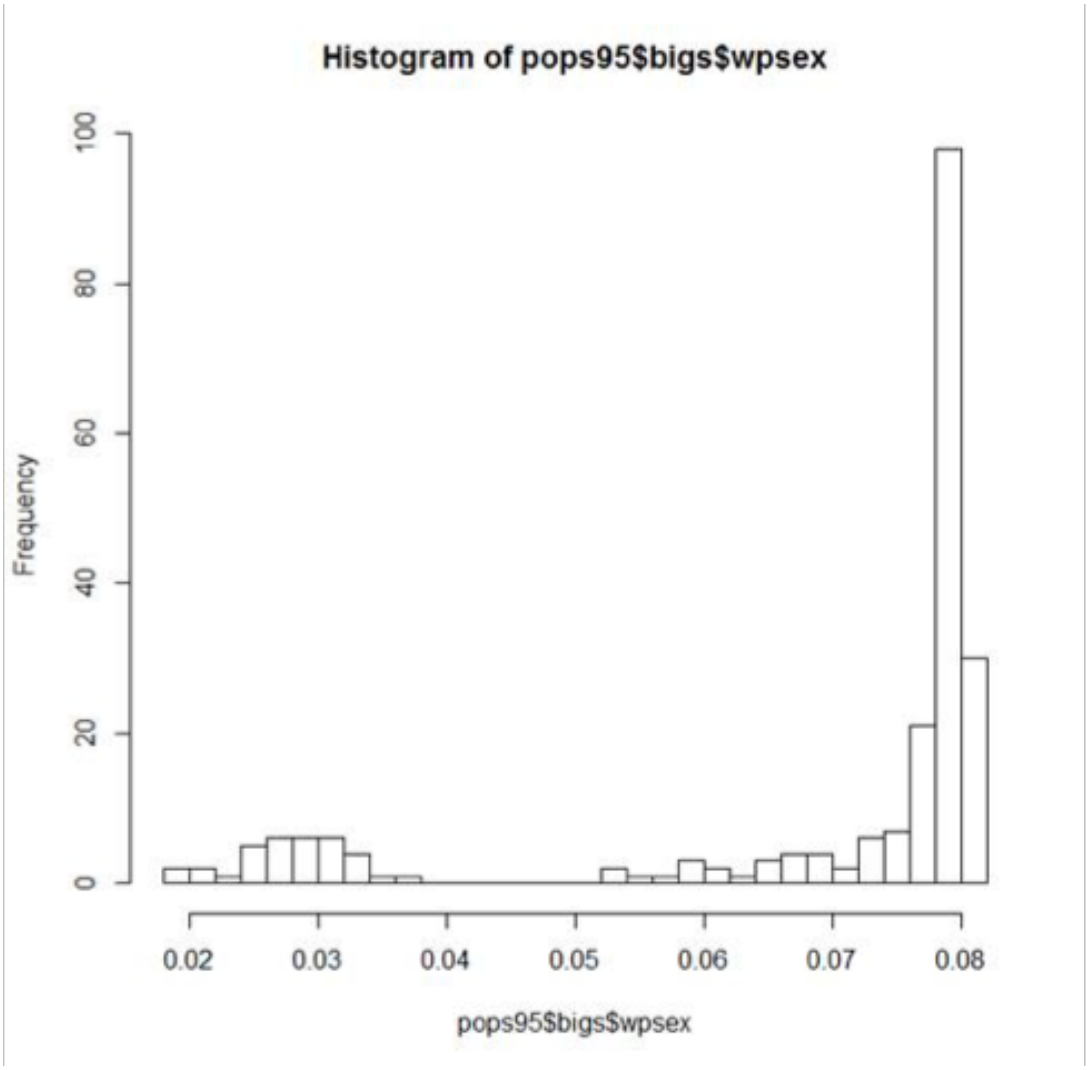
The microsatellites used in the first round of SBT CKMR were adequate for finding POPs using Mendelian-exclusion principles (see Appendix of Bravington et al., 2014). However, a lot of statistical processing was required to control false-positive rates to an acceptable level and demonstrate that false-negatives must be rare. As the new DArTcap genotyping (Bravington et al 2015) has been designed with the goal of identifying HSPs, which is much harder than finding POPs; finding POPs ought to be easier and clearer with the DArTcap genotyping results.

As in Bravington et al 2014, we again identified POPs using a classification statistic based on Mendelian-exclusion, but some changes to the method were required to deal with the new features of DArTcap data (see Appendix, Bravington et al 2017b). Figure 1 shows part of the histogram of the modified exclusion statistic, referred to as the Weighted-PSeudo-EXclusion (WPSEX) statistic, across all genotyped adult-juvenile pairs (about 66,000,000 comparisons). The POPs are visible as a small bump on the LHS. Most of the entire histogram (to the right hand side of the figure) has been left out, as otherwise the true POPs are too few to be visible compared to the very large peak of unrelated pairs. The peak of the unrelated pairs distribution is at 0.116, which is precisely where theory predicts it should be based on the allele frequencies of each locus.

The very large peak of unrelated pairs drops off very quickly to the left of ~0.08, and the fattish tail around 0.055 to 0.075 will contain a number of adult/juvenile HSPs or GGPs (Grandparent-Grandoff-spring Pairs), which should be somewhat rarer than true POPs on demographic grounds. The POPs are clearly separated from non-POPs. This separation is much more obvious with the new DArTcap data than it was for the original microsatellite data; demonstrating that the 1500 low-information SNP loci from DArTcap are performing better than 25 high-information microsatellite loci, at about roughly half the cost.

The results presented in Figure 1 only uses only adults from 2010 onwards and, hence, exclude the POPs already found via microsatellites. However, we also DArTcapped those particular pairs-of-samples already identified as POPs in original study, as a check, and they were all clearly identified as POPs using the new DArTcap method. Interestingly, we also processed one curious adult/juvenile pair from 2012, which was clearly not a POP according to microsatellites, but nevertheless remarkably close (just two unambiguous Mendelian exclusions in 25 loci compared). The DArTcap WPSEX statistic for this pair was around 0.06, consistent with being a Grandparent-Grandoff-spring Pair or HSP.

Figure 1: Identification of Parent-Offspring Pairs (POPs) via weighted-pseudo-exclusion (wpsex). Low values (left-hand side) indicate POPs. The x-axis has been truncated on the right-hand side to omit the majority of the very large peak peak of unrelated pairs, which would dwarf the small "bump" of POPs on the LHS if included.



The distribution of the POPs found in this study appear generally consistent with results from the previous study (Table 4). No POPs were found where the parent was caught in the same year as the off-spring was born. As noted above, such comparisons are excluded from the model anyway, to avoid potential bias. As in 2012, we also read the ages of all adults in POPs using otoliths collected in Indonesia at the same time as genetic sampling. The modal age of parents at time of off-spring's birth was 13 or 14 years. All bar one of the parents were 8yo or more at off-spring-birth, as was the case in the previous study. There was, however, one parent inferred to be 7 year-old: 14yo at capture, 7 years after its off-spring was born. Uncertainty of ±1year in otolith-derived ages is not uncommon, so this may be an age-estimation error. We are in the process of confirming that all parental otoliths do correspond to the genetic samples, as there is at least one case in the previous study where tissue and otolith samples must have been mixed.

Table 4. Distribution of Parent-Offspring Pairs by year juvenile birth year (rows) and adult Capture year (columns).

	2006	2007	2008	2009	2010	2011	2012	2013	2014
2002	0	0	0	0	0	0	0	0	0
2003	0	5	1	2	0	0	0	1	0
2004	0	2	0	0	3	0	0	0	0
2005	1	4	5	4	1	0	1	1	2
2006	0	4	3	2	0	0	0	0	0
2007	0	0	3	4	1	3	2	0	2
2008	0	0	0	0	0	1	1	1	0
2009	0	0	0	0	0	1	1	1	0
2010	0	0	0	0	0	3	1	4	0
2011	0	0	0	0	0	0	1	2	1
2012	0	0	0	0	0	0	0	1	1

The distribution of POPs in Bravington et al 2014 indicated that adults younger than 12 years may spawning every second year, i.e. exhibiting skip-spawning behaviour. Only two of the new POPs in the current study involved an adult caught at age 12 or less, but in both cases the number of years between off-spring birth and adult capture was even. This result adds additional weight to the hypothesis that younger adult SBT skip-spawn[4]. Note that the occurrence of some 4 year old individuals among the more recent juvenile samples will make this pattern harder to detect, should they continue to be present in appreciable quantities in future.

3.1.2 Half-Sibling Pairs

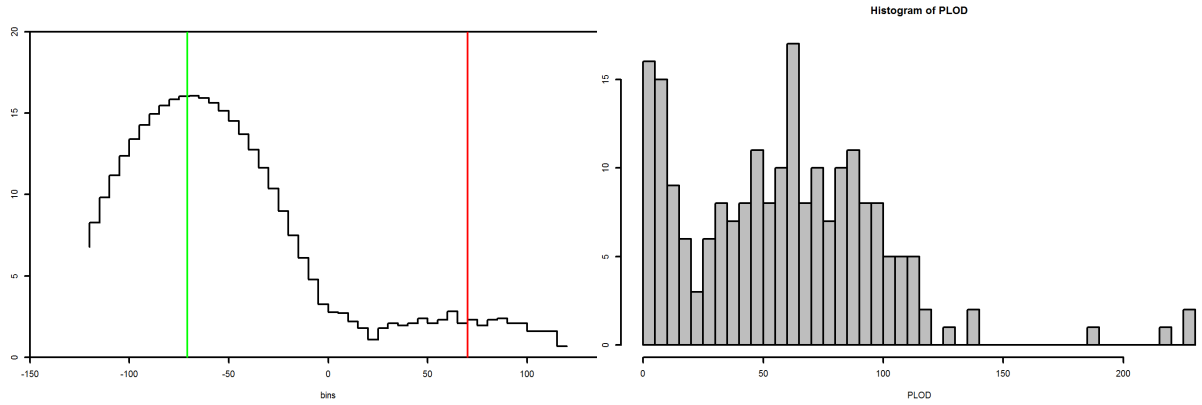
Among 10,809 juvenile genotypes, we found 140 definite HSPs and 4 Full Sibling Pairs (FSPs). The true number of HSPs is expected to be about 10% higher than 140, because of false-negatives that

[4] Note that the occurrence of some 4 year old individuals among the more recent juvenile samples will make this pattern harder to detect, should they continue to be present in appreciable quantities in future samples.

are inevitable (and expected) consequence of the statistical criteria used to ensure exclusion of all false-positives (see Bravington et al 2015). The HSPs and FSPs are quite clearly identified in Figure 2. The distributions of the PLOD test statistic match the predictions of genetic theory, indicating that the new genotyping and HSP-finding processes work reliably. The details of the genotyping and HSP-finding analyses are provided in the Appendix.

The proportion of HSPs where both individuals were caught in the same year is somewhat higher than would be expected under a completely random breeding scenario[5]. This is evidence of "lucky litters", i.e. variable survival between spawning events[6] within each year class, which is also the only way to explain the 4 FSPs identified[7]. However, SBT are clearly not a sweepstake reproduction species; the proportion of juveniles in same-cohort HSPs is still very small (<1%). This result confirms the conclusion of the original CKMR analysis (Bravington et al 2014) that it is a reasonable approximation to treat all POP comparisons as statistically independent. Note that within-cohort HSP comparisons are not used in our CKMR models; the HSP information comes entirely from cross-cohort comparisons.

Figure 2: Half-Sibling Pairs (HSP). Left hand side: log histogram to show all individual comparisons included in the analysis. The large dome on the left of zero is the Unrelated Pairs (UPs). The flatter, smaller dome on the right are the HSP and a few FSP. The green and red lines are theoretical means for the distribution of UPs and HSPs, respectively. Right hand side panel: is a histogram of PLODs above zero, which shows the HSP (PLOD ~35-130) and full-sibling pairs (PLODs above ~140).



[5] In a completely random breeding scenario, every juvenile sampled would have randomly selected its mother independently from the pool of potential mothers (weighted according to their relative fecundities), and likewise its father. The key word is independently (i.e. between juveniles in our sample). This does apply to juveniles in different cohorts, but the HSP data show that is not entirely true for juveniles in the same cohort.

[6] Each SBT on the breeding ground spawns on many nights per year. Post-fertilization larval survival rates may well differ between nights.

[7] The chance of a female breeding twice independently with the same male is inverse to adult abundance, so cross-cohort FSPs should be about a million times rarer than HSPs. The same applies to same-cohort FSPs, unless some spawning-events, where one female and a small number of courting males all release eggs and sperm together, have higher post-fertilization survival than others. Unsurprisingly, all 4 FSPs found were same-cohort.

3.1.3 Maternal/Paternal ratio

From analysis of mitochondrial DNA, the Maternal/Paternal proportion in the HSPs is close to 50:50 (i.e. whether the shared parent is the Mother or the Father), both for same-cohort and cross-cohort HSPs. The mean number of cohorts separating each HSP is very similar for Maternal vs Paternal HSPs, so that the SSB turnover rate must be similar for both sexes, something which is not biologically obvious in advance. This validates an assumption underlying the exploratory use of combined-maternal-and-paternal HSPs in the CCSBT operating models for the 2017 assessment (Hillary et al 2017b), where adult sexes are not distinguished.

3.2 Stand-alone CKMR assessment

3.2.1 Abundance and Spawning biomass- POP only versus POP+HSP models

The estimates of abundance from the new POP+HSP model and data are fairly similar to the values from the previous POP-only study (Bravington et al 2014). Figure 3 shows that the new estimates of SSB are about 10% higher on average— a degree of change which is consistent with that expected from sampling variability, given there were 45 POPs available in the original study, whereas there are 76 in total in the updated data series. The downwards trend in SSB also seems milder with the addition of the new data. This general pattern of new estimates being slightly higher, and trends apparently weaker, is repeated in most of the other abundance-related statistics.

It is noteworthy that there is less of a difference between the old and new estimates earlier in the series, i.e. in 2002 relative to 2010. CKMR data tend to contain more information in the earlier part of a time series than the more recent part, because there have been more opportunities to recapture parents of juveniles born in the early years. Hence, any inference made in in 2010 about abundance close to 2010 would have been particularly uncertain. Similarly, inferences about 2014 made now are less certain than those about 2010 made now.

All options suggest a small preponderance of females in the adult population, in the range 52–58%. The increase in fecundity-at-age for older (female) SBT appears less marked in the new data than was apparent in Bravington et al 2014 (Figure 4). The original estimates were of course based on just 20 female parents in total.

Overall, there are no obvious strong differences between results from the original SBT study (using POP data only, up to 2010) and the new model (using HSPs as well as POPs, and data through to 2014).

Figure 3: The original POP-only and new POP+HSP model estimates of adult biomass presented as total biomass of SBT aged 10 years and older. Black is the original model, which has been refitted using up-to-date software and modified code yielding very similar numbers to Bravington et al (2014); red is the new POP+HSP model with fixed selectivity; green is new POP+HSP model with estimated selectivity. Similar ratios are seen for numbers of SBT aged 10 years and older.

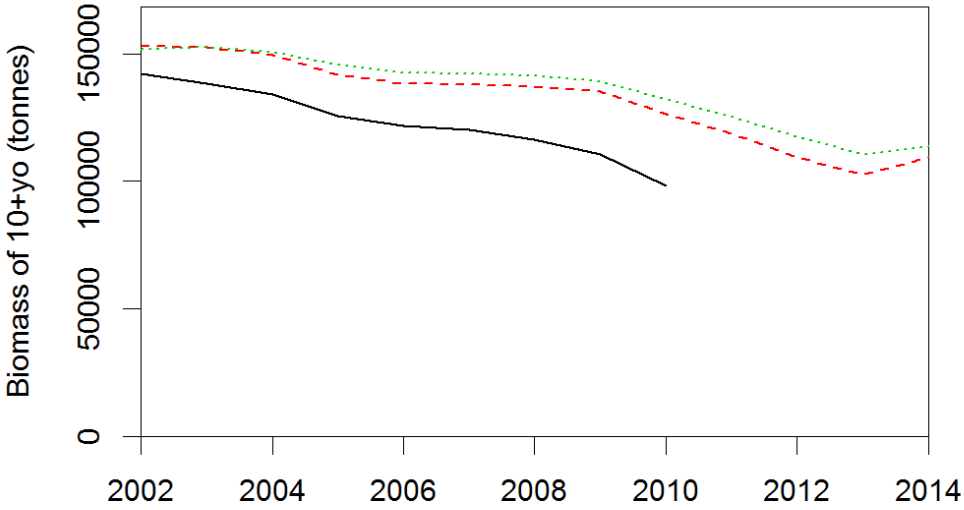
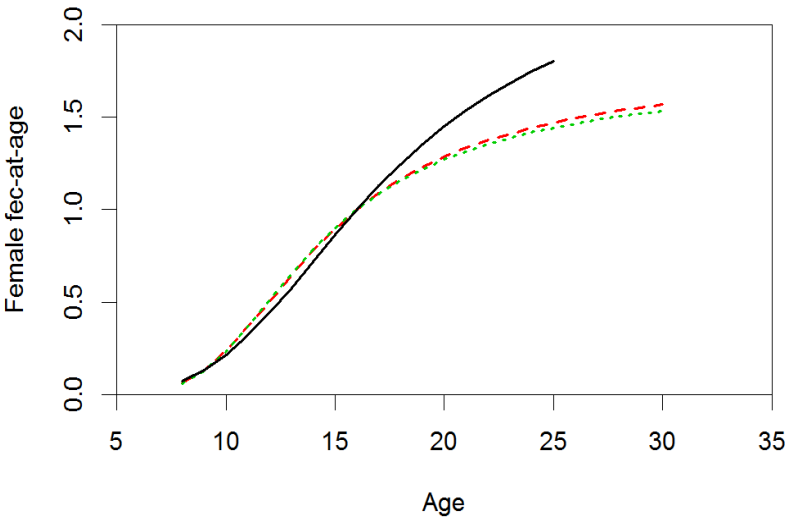


Figure 4: Female fecundity-at-age, normalised to 16 years old, from original POP-only model (Bravington et al 2014) and the new POP+HSP model and data. Black is original; red is new with fixed selectivity; green is new with estimated selectivity.

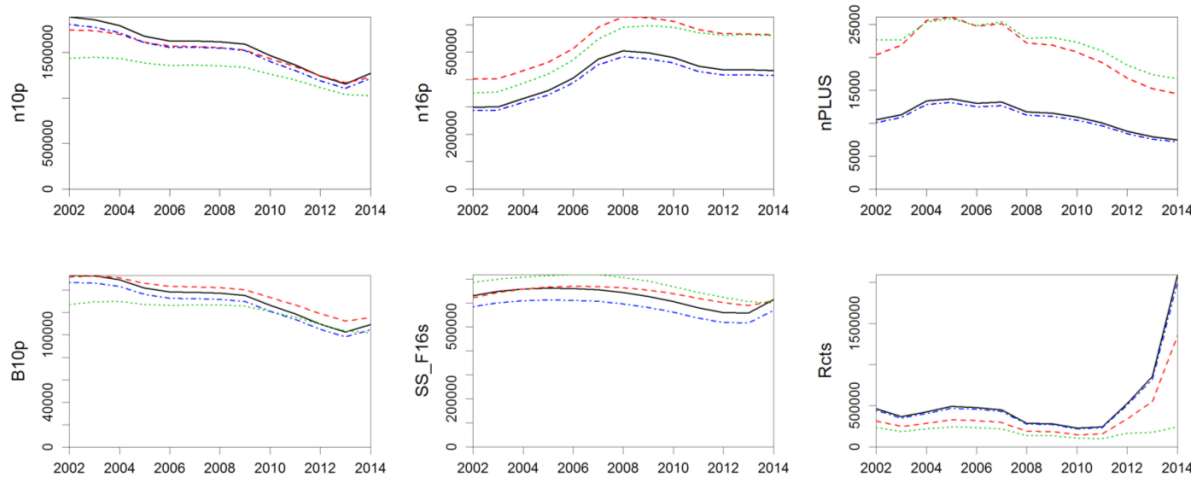


3.2.2 Influence of Alternative Model assumptions on abundance related estimates

Turning to the new model options, the overall summary statistics of biomass and numerical abundance change rather little across the model options explored (Figure 6), but there are differences in the age-specific components (n16p, nPLUS, Rcts). whereby the options with estimated, rather than fixed, selectivity predict more old and fewer young fish.

All model options show very strong incoming cohorts of 8yo (lower RHS panel) from about 2012 onwards, with the exception of the dotted-green option, where the LSfreq has been down-weighted by a factor of 100 so that it carries almost no information. Unsurprisingly, that option is slower to recognize incoming cohorts. By 2014, those cohorts have started to make an impact on overall SSB and TRO, so substantial upward trends in TRO and SSB would be expected from 2015 onwards as these recent adults grow bigger. These cohorts were born from 2004 onwards.

Figure 6: Abundance related estimates for six different components of the adult age-structure (each panel) from four different POP+HSP models (different lines). n10p = abundance of 10+ year olds; n16p = abundance of 16+ year olds, nPLUS = abundance of 30+ year olds; Rcts = abundance of 8 year olds recruiting to the adult component of the population; SS_F16s = Total Reproductive Output of females standardised to units of 16 year old female; B10p = biomass of 10+ year old. Lines reflect different models: black solid: fixed selectivity, with *αHSP* fixed at 1; blue dot-dash: as for black except with *αHSP* estimated; green dotted: as black except with LSfreq downweighted 100-fold; red dash: as black except with selectivity estimated.



3.2.3 Total Mortality and Selectivity

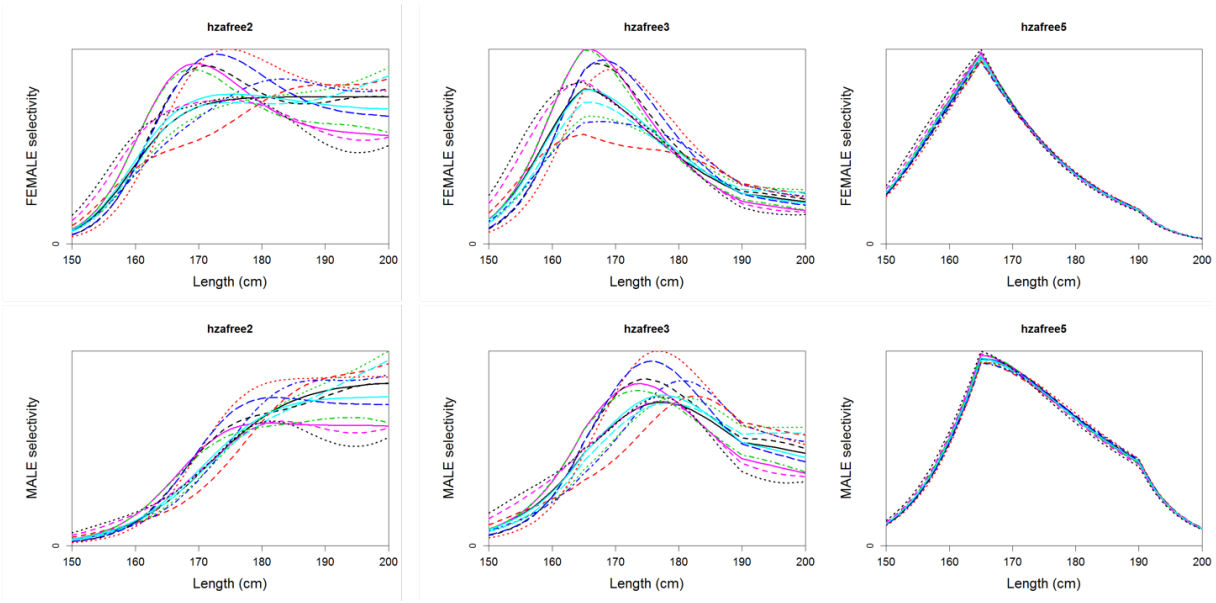
Results from exploring the influence of a wide range of model options different model options indicated the main differences in terms of selectivity and total mortality occur in the following areas:

- The estimated-selectivity options gives somewhat higher survival for younger adults (0.85 vs 0.80), and preference for dome-shaped selectivity. The estimated mean age of 8+yos in the

population is about 1 year greater for the estimated-selectivity model, with that gap increasing slightly over the time series (2002–2014).

- A selection of estimated selectivity curves are shown in Figure 7. The preference for a dome-shape selectivity persists even when the LSfreq is heavily down-weighted (right-hand column).
- When *αHSP* and selectivity are both free, rather than fixed, and when LSfreq is not downweighted, the resulting estimate of *αHSP* takes is nonsensical (α is estimated at ~2, which is more than 3 standard errors away from its maximum mathematically-plausible value of 1). For this option, the biomass estimates do change substantially, but as the parameter estimates make no biological or mathematical sense, there is no reason to consider this option plausible.

Figure 7: Annual selectivities-at-length with females on top row and males below. Left column are "fixed" selectivity (i.e. matching the residence time); middle column is "free" (estimated) selectivity; right column is free selectivity but with LSfreq downweighted by 100-fold. Different-coloured lines are for different years, showing annual variability. The abrupt kinks in the right-hand plots result from the piecewise-linear form used for base selectivity-at-length. These are smoothed out in the middle plots via the year-specific selectivity adjustments, but in the right-hand plots the LSfreq data is so heavily down-weighted that there is no incentive for the model to bother with much smoothing.



A few other options have been investigated (e.g. different z by sex; more flexible size-fecundity relationship for males), but without much obvious effect. Down-weighting the LSfreq data has interesting effects. The preference for dome-shaped selectivity remains, even though there is little impetus from the (down-weighted) data to drive it directly. The fit to the CKMR summary statistics is in fact slightly better with down-weighted LSfreq (e.g. less tension between HSPs and POPs when α HSP is fixed at 1), perhaps because the down-weighted options have more freedom to adjust other parameters, such as growth rates, when there is no need to match the LSfreq data closely.

Whether selectivity is fixed or free, the down-weighted options prefer higher survival rates for young adults (from 0.80 to 0.87 if selectivity is fixed; from 0.85 to 0.91 if selectivity is free). There is still remarkably little change in the main summary statistics (numbers of adults, SSB, TRO), but the higher survival rates do correspond to substantially different estimates of the absolute abundance of 8yo recruitments, and of overall age composition.

The internal CV on most biomass and numerical-abundance statistics is about 10–15%. It is systematically lower when *αHSP* is fixed, as, in this case, HSPs as well as POPs contribute directly to the *absolute* level of abundance, so the latter is being estimated based on 178 kin-pairs (HSPs and POPs) rather than on 76 (POPs alone). Although internal CVs are lower than in the original Bravington et al 2014 model— i.e., precision is better— the reduction is not as great as one might expect based purely on the increase in number-of-kin-pairs. The reasons for this are: (i) the new model is more flexible and more parameters are being estimated, and (ii) the LSfreq data is handled differently in the new model and its contribution will be less than in the original study, so it contributes less apparent precision. The range of point estimates from alternative model assumptions is now of the same order of magnitude as the reported CVs, and it appears (from evidence of some internal tension between the datasets; see below) that the LSfreq is still being weighted too highly. Thus the internal CVs must still be somewhat low, and more exploration is warranted. This was less obvious in 2013, perhaps because fewer options could be explored, and the CVs were larger anyway because of reduced sample size.

3.2.4 Model fits and Diagnostics

Aside from one case:

- All model options give reasonable to good fits to a range of close-kin summary statistics (see section below). So far, the best fits to CKMR statistics have been obtained with fixed selectivity; free-selectivity models do show some tension between POPs and HSPs, presumably reflecting tensions between CKMR and LSfreq data (or the handling of the latter in the model).
- All options fit equally badly to the sex-ratio in the Indonesian LSfreq data, in particular showing less trend (towards more males) than is seen in the observed data;
- The fits are compatible with *αHSP* =1 , which would in fact be the point estimate if the two sexes were constrained to have the same value of *αHSP* . This is a “no surprises” result; the representation of individual fecundity appears consistent with the CKMR data, and the POPs and HSPs are telling the same story. Given the current totals of about 75 POPs and 140 HSPs, the lower confidence interval on *αHSP* (which is set by the ratio of those totals) is about 0.85; thus, values slightly below 1 are not ruled out, but there is no evidence for them in the observations. In particular, there is no evidence for “hidden population structure”, whereby some adults persistently tend to have offspring that go to South Africa rather than the GAB. CKMR alone can never exclude the possibility that some juveniles do go to South Africa, rather than to the GAB, but there is no evidence that their siblings are also more likely to do so.

- When selectivity is estimated and *αHSP* =1 , there is some mismatch between the POP and HSP totals— not statistically significant, but nevertheless noticeable. Letting *αHSP* be estimated would not alleviate matters, since the preference of the model would be to make α ˆ HSP >1 which is not possible (see below). In other words, the best *practical* fits with estimated selectivity still have *αHSP* =1 .

The only model option checked so far that does *not* give a reasonable fit (at least to the CKMR data) is when both *αHSP* and selectivity are estimated, rather than fixed. This is a somewhat paradoxical result— the worst fit coming from the option with the most flexibility in fitting.

Similar results were obtained from the POP+HSP OM in Hillary et al (2017c) (bearing in mind that the OM is actually formulated in terms of a di erent parameter, qhsp = 1/αHSP; the OM results are here reinterpreted in terms of αHSP). That OM also preferred *αHSP* somewhat above 1, but again the results were consistent with the maximum plausible, and most desirable, case of αHSP = 1. Since there is no strong *a priori* reason to expect αHSP < 1, it is merely a possibility that needed to be allowed for, and also no actual evidence that it is below 1 despite reasonable sample sizes (numbers of POPs and HSPs), it seems fair to assume a "base case" of αHSP = 1 at least until there is any clear evidence (from future CKMR analyses) to the contrary.

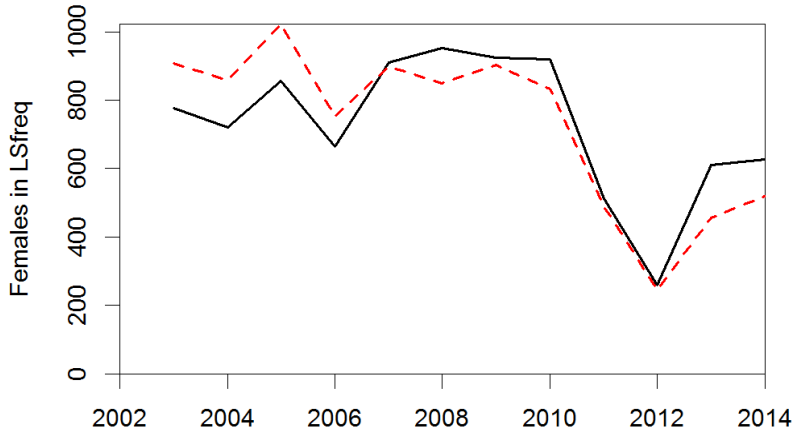
Fits to A@LS data are good across options (Figure 8). Fits to the length-frequencies LSfreq look very impressive for all options (not shown). Unfortunately, the plots actually contain no useful diagnostic information, because the annual selectivity adjustments adjust automatically to match the observed data quite closely. In fact, there must be substantial differences in how hard the annual adjustments are having to work, leading to statistically better fits for certain options, but these are not easy to see; some better way of plotting the diagnostics is needed here. Fits to sex frequency in LSfreq, however, are strikingly bad for all options despite the annual selectivity adjustments, which so far are not sex-based (Figure 9).

3.2.5 CKMR diagnostics

The match between observed and predicted CKMR summary statistics seems generally good. Table 5 shows some simple summaries. Predictions within ± 1 SE of the observation are entirely consistent with the data, and between about 1 and 2SE are just about consistent; above two SE indicates misfit. Fits to males are worse than to females (and less effort has been spent trying to represent male fecundity in the model). In all options, male turnover rates (corresponding to *Pdt* in the Table) are predicted somewhat too low, and male parental length is predicted significantly too high. This latter difference is mitigated when the *LSfreq* data is downweighted. Other than that, the only noteworthy feature of these summaries is the moderate mismatch between HSP and POP totals in option *Bc* (free selectivity, full weight to LSfreq).

Figure 8: Fits to A@LS data, for females and males. Thick black lines are observed mean ages; coloured and broken lines are predictions from different options. The three groups of lines are: top, for big fish i.e. with lengths in the upper tercile (third) of all sampled lengths; middle, all fish; bottom, for small fish whose lengths are in the lower tercile. 2002 is missing because of concerns about reliability of sex measurements in 2000–2002. 2012 is missing because age estimates from otolith-reading are not available for that year.

Figure 9: Fits to total number of females in LSfreq data. Thick black line is the observed count; broken pink line is the prediction (allowing for total sample size that year). All options give essentially the same prediction, so only one is shown. Absolute numbers have been plotted to show the very considerable magnitude of the discrepancy (⩾ 6 Standard Deviations in many years).



The one badly-fitting option

The one really bad fit to CKMR data occurs when selectivity and α HSP are both estimated. In theory, this should give the best fits because it has the most flexibility to adjust parameters, but in practice it gives the worst! The problem is that the point estimate of α HSP rises to about 2, for both sexes; this about 3 standard deviations above the maximum plausible value of 1, so clearly is

not just an unlucky accident of data. This option leads to only slight improvements in the fit to the observed CKMR statistics (α HSP being a parameter not a observed statistic), so presumably the real gain is somehow in an improved fit to LSfreq. Unsurprisingly, such a high α does markedly change the abundance estimates, but since the parameter estimates makes no sense, this result can be disregarded . Nevertheless, it does indicate some unresolved tension between datasets.

It seems likely that selectivity in the LSfreq data is the root cause of misfit; some more work is needed here. The LSfreq data has very large sample sizes, and unless carefully handled it tends to bully the other datasets to get its own way in the log-likelihood; that is, the model ends up willing to sacrifice appreciable goodness-of-fit to CKMR and A@LS data in order to accommodate minor nuances of the LSfreq data (although the models checked so far are still unable to match the sex proportions in the LSfreq, as mentioned). In principle, this dominance should be eliminated by the annual selectivity adjustments, but evidently our attempts at that have not yet met with complete success. There are two likely reasons:

- the basic model for selectivity and how it varies might not yet be formulated appropriately, especially with respect to sex;
- the variance of the year-to-year fluctuations in selectivity has been pre-estimated (see Appendix) and fed into the entire model as a fixed parameter. This would be OK if the pre-estimated variance was appropriate, but in fact it seems that rather more variance is required to match the data, especially with respect to sex frequencies. In particular, the pre-estimated variance was based only on year-to-year changes in length-frequency, not sex-frequency. The statistical software used to fit the model currently has the limitation that that only one random-effect variance parameter can be estimated inside the model; any other random-effect variances must be supplied a priori as fixed values. Currently, that one variance has been reserved for recruitment variability, so the annual selectivity variance has to be pre-estimated. It is preferable, at least in theory, to estimate all random-effect variances inside the model, and future versions of the software will eventually allow this.

Table 5 CKMR summary statistics: observed and expected, by sex. All options in this table have *αHSP* fixed at 1. *A/B*: fixed/free selectivity *c/d*: full weight/downweight LSfreq. *Hdt*: interval between birth-year in cross-cohort HSPs. *Htot*: total number of HSPs. *Pdt*: interval between offspring birth and parental recapture. *Pl*: length of parents at capture. *Ptot*: total number of POPs *dSE*: number of standard errors (computed empirically from observations) between prediction and observation.

	Hdt				Htot				Pdt				Pl				Ptot			
	F	dSE	M	dSE	F	dSE	M	dSE	F	dSE	M	dSE	F	dSE	M	dSE	F	dSE	M	dSE
Obs	3.07		3.05		45		57		3.56		3.33		171		175		36		40	
Ac	2.92	-0.5	2.96	-0.4	47.7	0.4	57.4	0.1	3.65	0.2	3.73	1.4	173	1.4	179	4.4	33.4	-0.4	39.7	0
Bc	3.04	-0.1	3.05	0	53.4	1.3	65.3	1.1	3.58	0.1	3.68	1.2	173	1.7	180	5.1	27.6	-1.4	31.7	-1.3
Bd	3.23	0.5	3.22	0.7	50.5	0.8	58.4	0.2	3.88	0.8	3.93	2.1	171	0	176	1.4	30.5	-0.9	38.6	-0.2
Ad	3.14	0.2	3.11	0.2	47.8	0.4	55.9	-0.1	3.84	0.7	3.88	1.9	171	0.2	177	1.8	33.2	-0.5	41.1	0.2

4 Conclusions

The main summary statistics of estimated adult biomass and numerical abundance seem fairly stable across a range of model options. The new POP data alone are consistent with what would have been expected from the model used in 2012. The new HSP data also seem to be consistent with the POP data and with a “no surprises” model of SBT reproductive biology, and with the assumptions made (but not possible to test) in 2012.

A key difference from 2012 study (regardless of whether selectivity is estimated or assumed to follow the 2012 assumptions) is that the HSPs are now providing a direct signal on overall adult *z*, and this seems broadly consistent with the overall *z* that was inferred under the assumptions in the original study. However, the new model does show some preference for a somewhat higher survival for young adults, and an overall dome-shaped selectivity. This difference would have some effect on turnover rates, and estimated incoming 8yo recruitments. What is not yet clear is how seriously to take that dome-shape. Since the treatment of selectivity for the *LSfreq* data (whether estimated or fixed at the 2012) is still not fully satisfactory, especially with respect to the observed sex ratios. Some further investigation is warranted to examine this in more detail.

Aside from the implications for SBT, this project has important implications for other CKMR studies. It has been the first opportunity to consider the combination of POPs and HSPs in the same model, and to reflect on the benefits of having both. This is a research area that is still unfolding— the focus in this project has been on finding the HSPs reliably, and on working out how to build models specifically for SBT— but some general principles have emerged. Our overall message is that both POPs and HSPs are necessary for standalone CKMR models of teleosts in general; neither is sufficient on their own. Simplifications are possible for life-histories where, for at least one sex, there is little change in fecundity through adulthood and between different adults. That applies to many mammals and some sharks— but few, if any, teleosts. The requirement for HSPs and POPs has implications partly for the design of sampling strategies for CKMR, but especially for the approach to genotyping. Finding POPs is comparatively easy and there are several genotyping techniques that might be adequate and affordable, but finding HSPs is certainly not easy and requires very high reliability from the genotyping method. The need to both keep costs down and reliability sufficiently high, places tight constraints on the genotyping technique. Although the approach that we have used may not be the only viable one, we have yet to see clear evidence that any other current approach will work. Any proposal for a CKMR study needs to justify quite carefully why its proposed genotyping technique will be adequate to reliably identify HSP.

There is an informal argument for why POPs as well as HSPs are necessary for CKMR in teleosts (and why, together, they are also sufficient). Stripped of details, such as, time-trends and multi-year sampling, which are important in practice for fitting to data, but confusing for a general picture, the argument runs as follows.

POPs lead directly to estimates of two things:

1. The pattern of fecundity with age (or size)— based on the different rates at which POPs are found per comparison with big/old fish, vs per comparison with small/young fish of the same sex.
2. How many "equivalent adults" are contributing to breeding— "equivalent", that is, in terms of successful reproductive output— based on the average per-comparison rate of finding POPs. That could be measured on any arbitrary scale, such as "equivalent 15yo fish" or "equivalent 175cm fish". Typically, this is only deemed to be of interest for females, since males are assumed to be in surplus.

The underlying biological model behind those estimates is very simple and transparent, and there is no serious room for ambiguity or model-dependence. Given a decent sample size of identified POPs, the precision should be quite tight. What POPs do not directly reveal, though, is whether the given number of equivalent adults comes from, for example, a lot of small, young adults, or from fewer big old ones. POPs also provide no intrinsic information on rate-of-turnover.

Those two pieces are essential parts of any useful stock assessment. Hence, to construct the original stand-alone CKMR model for SBT from POP data alone, it was necessary to make some reasonable but, at the time, untestable assumptions, and to take advantage of some special features of SBT biology and adult sampling. We have not yet seen those special features in any other species where CKMR has been proposed.

Half-sibling pairs (from different cohorts) also directly lead to two estimates, but of rather different things:

1. The rate of turnover of adult SSB (or, more accurately, of Total Reproductive Output)— based on average intervals between birth-years of HSPs.
2. Provided that αHSP =1, the number of "equivalent squared adults" that are contributing to breeding, again on an arbitrary scale, such as, "equivalent to the average squared fecundity of a 15yo fish". As with POPs, this is based on the average per-comparison rate of finding HSPs, but the complication with HSPs is that big adults are systematically over-represented. Not only are they more likely to be "tagged" through having more offspring among the first cohort of juveniles, they are also more likely to be "recaptured" through having more offspring among the second cohort too. Even with deliberate exclusion of same-cohort HSPs (where random as well as systematic variability comes into play), any systematic individual-level variation in fecundity across a few years will have this effect.

While HSPs on their own are not much use for teleosts— the number of "equivalent squared adults" does not relate to management in the direct way that, say, TRO or SSB does from POPs does— the combination of POPs and HSPs is very powerful. Under certain assumptions about the form of fecundity-size relationships, there is only one mean age of adults that will be consistent with both a given number-of-equivalent-adults and a number-of-equivalent-squared-adults. Once that mean age is established, under certain assumptions, the rate-of-SSB-turnover from HSPs leads to an estimate of average adult z (and natural mortality, if catches are known). This fills the gaps left by POPs alone, and allows at least a rough stand-alone CKMR stock assessment (rough, in that

only one aggregate z can be inferred; age-specific effects could not be estimated directly). No length- or age-composition data is required in such a case, though age and size data must be available for the genotyped samples. In practice, there would need to be enough general information on age-and-length to give a satisfactory model for individual-specific growth, but this can potentially be obtained from sub-sampled otolith data alone (A@LS), which is fishery-independent and not subject to selectivity. Note also that it is necessary to assume αHSP =1 in order to perform this rather potent piece of magic.

It is not yet clear how much, if any, fishery-derived length/age composition (LSfreq) data is required in practice for CKMR for teleosts in general. The SBT model remains statistically identifiable just from CKMR and *A@LS* data, even when the *LSfreq* data is diluted to very low levels. That, plus the conceptual argument just presented, suggest that technically there may be no need for *LSfreq* in order to fit a reasonable model. Such an option is appealing, in that it completely bypasses any potential problems caused by selectivity and, in particular, by changing selectivity. However, the SBT option with downweighted *LSfreq* is also rather unresponsive to incoming cohorts, which will take a long time to grow to a size where their impact on the TRO is evident in the CKMR data alone (the *A@LS* data responds faster, but not as fast as the *LSfreq* data, because the sample sizes are not that large). Consequently, the robustness of a completely *LSfreq*-free CKMR method needs to be balanced against the greater precision and responsiveness of a method that attempts to handle selectivity and fishery-derived length-composition— provided that the latter method actually works. This tradeoff ought to be considered not just in statistical terms, but also in the context of management needs. Now that there is more experience and confidence in the "mechanics" of CKMR, the whole question of how to use CKMR in Management Procedures is ripe for exploration.

5 References

Anon., 2017b. Report of the OMMP Technical Webinar on incorporation of half-sibling pairs in the CCSBT OMs for the 2017 update of stock status. CCSBT-ESC/0817/36.

Bravington, M (2014). Close-Kin Mark-Recapture for SBT: options for the longer term. Report for Scientific Committee CCSBT-ESC/1409/44. Commission for the Conservation of Southern Bluefin Tuna.

Bravington, M (2017). SBT kin-finding and genotyping update. CCSBT-OMMP/1706/12. Commission for the Conservation of Southern Bluefin Tuna.

Bravington, M.V., Grewe, P.M., and Davies, C.R. (2016a). Absolute abundance of southern bluefin tuna estimated by close-kin mark-recapture. Nature Communications 7:13162. https://doi.org:10.1038/ncomms13162

Bravington, MV, HJ Skaug, and EC Anderson (2016b). Close-Kin Mark-Recapture. In: Statistical Science.

Bravington, M, P Grewe, and C Davies (2012). Report of the Close-Kin Project: estimating the absolute spawning stock size of SBT using genetics. Scientific Committee Report CCSBT-SC/1208/19. Commission for the Conservation of Southern Blue n Tuna.

Bravington, M, P Grewe, and C Davies (2012). Report of the Close-Kin Project: estimating the absolute spawning stock size of SBT using genetics. Scientific Committee Report CCSBT-SC/1208/19. Commission for the Conservation of Southern Bluefin Tuna.

Bravington, M and C Davies (2013). Close-kin for SBT: where to now? Working paper to the Extended Scientific Committtee of the CCSBT CCSBT-ESC/1309/17. Commission for the Conservation of Southern Bluefin Tuna.

Bravington, M, P Grewe, and C Davies (2014). Fishery-independent estimate of spawning biomass of Southern Bluefin Tuna through identification of close-kin using genetic markers. FRDC Report 2007/034. CSIRO, Australia.

Bravington, M, J Eveson, P Grewe, and C Davies (2015). SBT Close-Kin Mark-Recapture: options for the medium term. Working paper to Extended Scientific Committee of the CCSBT, CCSBT-ESC/1509/19. Commission for the Conservation of Southern Bluefin Tuna.

Bravington, MV, HJ Skaug, and EC Anderson (2016). Close-Kin Mark-Recapture. Statistical Science. Volume 31, Number 2 (2016), 259-274.

Bravington, M, J Eveson, P Grewe, and C Davies (2017). SBT Close-Kin Mark-Recapture with Parent-Offspring and Half-Sibling Pairs: update on genotyping, kin-finding and model development. Working paper to Scientific Committee of the CCSBT, CCSBT-ESC/1709/12. Commission for the Conservation of Southern Bluefin Tuna.

Bravington, M, R Thomson, and C Davies (2017). Genotyping issues for CKMR on Atlantic Blue n Tuna. Atlantic-wide research programme on Bluefin Tuna (ICCAT GBYP PHASE 5) GBYP xxx/2016. International Commission for the Conservation of Atlantic Tuna.

CCSBT ESC (2015). Report of the Twentieth Meeting of the Scienti c Committee. Commission for the Conservation of Southern Bluefin Tuna.

Davies, C.R., Bravington and Thomson, R. 2017. *Advice on Close-Kin Mark-Recapture for estimating abundance of eastern Atlantic bluefin tuna: a scoping study.* Final Report to the ATLANTIC-WIDE RESEARCH PROGRAMME ON BLUEFIN TUNA (ICCAT GBYP – PHASE 5) of the international Commission for the Conservation of Atlantic Tuna. 34p.

Davies, C.R., Preece, A. and Hillary R. (2016). Initial consideration of forms of candidate management procedures for SBT. CCSBT Extended Scientific committee for the 21st meeting of the scientific committee; September 2016; Kaohsiung, Taiwan. CCSBT; 2016. ESC BGD-05.

Farley, Jessica; Davis, Tim; Bravington, Mark; Andamari, Retno; Davies, Campbell 2015. Spawning dynamics and size related trends in reproductive parameters of southern bluefin tuna, *Thunnus maccoyii*. PLoS ONE. 2015; 10(5):e0125744. https://doi.org/10.1371/journal.pone.0125744

Farley, Jessica; Eveson, Paige; Davis, Tim; Andamari, Retno; Proctor, Craig; Nugraha, Budi; et al 2014. Demographic structure, sex ratio and growth rates of southern bluefin tuna (*Thunnus maccoyii*) on the spawning ground. PLoS ONE. 2014; 9(5 e96392):15 p. https://doi.org/10.1371/journal.pone.0096392

Jessica Farley, Ririk Sulistyaningsih, Craig Proctor, Peter Grewe, Campbell Davies 2017. distribution of SBT in the Indonesian longline catch and close-kin tissue sampling and processing. Working paper prepared for the Extended Scientific Committee for the Twenty Second Meeting of the Scientific Committee, Yogyakarta, Indonesia, 28 August -2 September, 2017. CCSBT-ESC/1709/9.

Hillary, R, A Preece, and C Davies (2013). Updates to the CCSBT operating model including new data sources, data weighting and re-sampling of the grid. Tech. rep. CCSBT-ESC/1309/15. Commission for the Conservation of Southern Bluefin Tuna.

Hillary, Rich; Preece, Ann; Davies, Campbell; Kurota, Hiro; Sakai, Osamu; Itoh, Tomoyuki; et al. A Scientific alternative to moratoria for rebuilding depleted international tuna stocks. Fish and Fisheries. 2016; 17(2 N/A):469-482. https://doi.org/10.1111/faf.12121

Hillary, R, A Preece, and C Davies (2017a). Updates required for new data sources and reconditioning of the CCSBT OM. Tech. rep. CCSBT-OMMP/1706/4. Commission for the Conservation of Southern Bluefin Tuna.

Hillary, R, A Preece, and C Davies (2017b). Summary of initial incorporation of the Half-Sibling Pair (HSP) data in the CCSBT Operating Model. Working Paper to CCSBT-OMMP Webinar, 21 July 2017. Commission for the Conservation of Southern Bluefin Tuna.

R. M. Hillary, A. L. Preece, and C. R. Davies, N. Takahashi, Japan O. Sakai and T. Itoh (2017c). Reconditioning of the CCSBT Operating Model in 2017. Working paper prepared for the Extended Scientific Committee for the Twenty Second Meeting of the Scientific Committee, Yogyakarta, Indonesia, 28 August -2 September, 2017. CCSBT-ESC/1709/14.

R. M. Hillary, M. V. Bravington, T. A. Patterson, P. Grewe, R. Bradford, P. Feutry, R. Gunasekera, V. Peddemors, J. Werry, M. P. Francis, C. A. J. Duffy and B. D. Bruce (2018a). Genetic relatedness reveals total population size of white sharks in eastern Australia and New Zealand. Scientific Reports Vol 8, Article number: 2661. DOI:10.1038/s41598-018-20593-w

Hillary, Rich; Preece, Ann; Davies, Campbell, 2018b. Data generation & changes to SBT OM. Working paper prepared for the 9th meeting of the Operating Model and Management Procedure Technical Group, Seattle, Washington, 18-22 June 2018. CCSBT-OMMP/1806/04.CCSBT-OMMP/1806/04.

Hillary, Rich, Preece, Ann, Davies, Campbell, 2018c. Initial MP structure and performance. Working paper prepared for the 9th meeting of the Operating Model and Management Procedure Technical Group, Seattle, Washington, 18-22 June 2018. CCSBT-OMMP/1806/05(Rev.1).

6 Appendix

APPENDIX: MODEL DETAILS FOR SBT CKMR 2018

Parameters are estimated within a REML framework, using Laplace Approximation and automatic differentiation to approximate the marginal log-likelihood allowing for random-effects. This is the same statistical framework used in e.g. TMB and ADMB-RE (ref** Skaug&Fo), but implemented via our in-house ADT software, which is far easier to debug. There is an outer layer of 3000 lines of R (ref**) code to organize the data, set up the model options, oversee parameter estimation, and extract summary statistics. The inner code, which implements the population dynamics and computes the log-likelihood from the observed data, consists of about 1500 lines of Pascal code (similar to C), about half of which is "housekeeping". Efficient derivative computation is crucial for Laplace Approximation, and this is achieved by automatic differentiation using the source-code transformation tool Tapenade (ref**); our in-house ADT software organizes the calls to Tapenade. Runtime to fit a model is 5–10 minutes, and no numerical problems have been seen during fitting.

Particular model options, e.g. which covariates are allowed to affect which parameters, are implemented through R's standard syntax for statistical formulae, so it is possible to explore many different options without touching the underlying Pascal code. For example, most of the options follow the current OM practice in assuming that survival probability[1] should be independent of sex and year, constant between ages 8 and 24, and then following a linear trend (actually on a logistic scale) to meet the plus-group z at age 30. This can be specified like so:

```
logit_psurv_formula = ~ I( pmax( age, 24)- AMAX) - 1
```

To allow dependence on sex, the formula would become:

```
~I(pmax(age, 24) - AMAX) %in% sex - 1
```

In the descriptions below, subscripts are used to indicate which covariates are *potentially* allowed to affect which variables. In most options, only a subset of the potential covariates are allowed, and this is noted.

The rest of this Appendix presents details of the population dynamics, and treatments of the four datasets, following a table of notation. For clarity, the words "we assume that..." are mostly omitted, but should be read implicitly throughout; e.g., the text *states* that growth follows individual-specific von Bertalanffy curves, but what is meant is that we *assume* that to be the case and have coded the model accordingly.

[1] In fact, the formula applies to the transformed mortality rate, $\text{logit}\left(e^{-z}\right)$, and there is an implicit offset of $\text{logit}\left(\exp\left(-z_{A_{\max}}\right)\right)$, the plus group survival, discussed below.

TABLE 1. Main notation. Many of these quantities are always disaggregated in practice by subscripts, e.g. to denote age.

Symbol	Meaning	Notes
y	Year; year-of-capture	Adults only
s	Sex	Adults only
ℓ	Length; length-at-capture	Adults only
b	Year of birth	Juveniles only
n	Numbers in population	
ϕ	Annual fecundity	Relative to the fecundity of a 170cm fish of that sex
σ	Selectivity	Chance of occurring in the LSfreq data, relative to a 170cm fish of that sex
z	Mortality rate	Annual survival probability=$\exp(-z)$
$A_{\max}$	Plus-group age	No growth and no age-specific changes in survival after that age
$\bar{a}_{\max sy}$	Mean age within plus-group	
$m_{\mathrm{P}},\ m_{\mathrm{H}}$	Number of pairwise comparisons	m_{P} relates to POPs; m_{H} relates to HSPs.
i, j	Labels for individual fish involved in a pairwise comparison	
K_{ij}	Measured kinship of fish i and j	If i is adult and j is juvenile, then the possible values are POP or UP. If i and j are both juvenile, the possible values are definite MHSP, definite PHSP, or UP/false-negative-HSP.
#HSP	Number of HSPs found	
$\mathbb{E},\mathbb{V},\mathbb{C},\mathbb{I}$	Expectation (mean); variance; covariance; indicator function	
$\triangleq$	"is defined as"	For temporary variables unworthy of inclusion in this table

(1) The notation $\mathbb{P}[X|\{y\}\,z]$, for conditional probability of some event X given covariates y and z, means that covariate y is formally required in the conditioning because of the previous application of a probability manipulation such as Bayes' theorem, but in practice is irrelevant to this particular probability.

(2) "Prime" variables, such as a' for age, are used in summations to distinguish between separate occurrences of the same *type* of variable in a single formula.

1. Population dynamics

Population dynamics follows a standard sex- and age-structure with annual time steps, using mortality rates but no explicit catches:

$$n_{s,y+1,a+1} = n_{sya} \exp\left(-z_{sya}\right);\ a < A_{\max}$$
$$n_{s,y+1,A_{\max}} = \left(n_{s,y,A_{\max}} + n_{s,y,A_{\max}-1}\right) \exp\left(-z_{sy,A_{\max}-1}\right)$$

which automatically imposes the constraint that plus-group survival is the same as in the preceding age-class. No year-dependent options have actually been tried yet. Plus-group age $A_{\max}$ was set to 30, except when repeating the 2013 analysis where 25 was used; data-wise, all measured ages above $A_{\max}$ were treated as "30 or more".

For CKMR probabilities, it is necessary to keep track of *some* information about ages within the plus-group. The mean age within the plus-group evolves according to :

$$\bar{a}_{\max s,y+1} = \frac{\left(\bar{a}_{\max sy} + 1\right) \times n_{syA_{\max}} + A_{\max} \times n_{sy,A_{\max}-1}}{n_{sy,A_{\max}} + n_{sy,A_{\max}-1}}$$

and we assume that age-within-plus-group follows approximately an exponential distribution with that mean.

The 2013 analysis, without the benefit of HSPs, *required* the assumption of equal male and female survival rates by age. With HSPs, it is at least possible to relax that assumption. In limited testing so far, though, letting survival rate depend on sex has not made much overall difference, so all the options reported in the main paper have assumed equal male and female rates.

Minimum age is 8— i.e., animals enter the model only when they reach age 8, the first age with evidence of successful breeding.

"Recruitment" at age 8 is log-normally distributed around a constant mean. The ratio of of males-to-females at age 8 in the population is an estimated parameter, but it is assumed constant over time.

The CV of n_8 was estimated at 0.35 using 2002–2010 data (as in 2013 model). With 2011–2014 data added— a period which include some very strong 8yo recruitments— the estimated CV rises to 0.41. Different model options do seem to have much effect on the estimated CV of 8yo recruitment.

1.1. **Plus-group survival rates.**

Catch-curve analysis of all post-1995 otoliths above age 30— an age above which growth has slowed enough that length-based selectivity should not affect the sampled age composition— shows an appreciably higher slope (Eveson**?). These data were fitted in a preliminary analysis by a Poisson GLM with parameters β and z such that

$\log \mathbb{E}\left[N^{\text{oto}}_{sa}\right] = \beta^{\text{oto}}_{0s} - z_{s,A_{\max}} \times (a - 30);\ a \geqslant 30$

There is a suggestion from this analysis, albeit not statistically significant, that plus-group survival rates are lower in males than females.

Except when repeating the 2013 analysis, the estimate $\hat{z}_{s,A_{\max}}$ and its variance (by sex, if appropriate) from the preliminary analysis were incorporated as offsets in the main CKMR model, with the difference $z_{s,A_{\max}} - \hat{z}_{s,A_{\max}}$ treated as a random effect of known variance. In the 2013 model (and its reincarnation

here), survival rate from age 25 up was estimated as a single parameter, and the composition of ages within the 25-and-up category was not used at all.

1.2. **Initial age composition.** In the first year of the model ($y_1 = 2002$), numbers-at-age are by default assumed to follow on average the catch-curve with slope ... (corresponding to constant average 8yo-recruitment in the past). These expected numbers in each age-class are then modified by cohort-specific random effects, just like incoming 8yo-recruitments in subsequent years. This is the same formulation as in 2013.

There is an option to estimate extra parameters for initial numbers-at-age, to allow general dependencies of the form $\log n_{s,y_1,a}|s,a$. This would accommodate, for example, changing pre-adult exploitation rates in the years before 2001, or (somehow) different sex frequencies in the initial spawning stock in 2002. Such options have not yet been explored, but might help deal with the mismatch to observed sex frequencies noted in the main text; nevertheless, it would be nice to have some insight into why such differences might have arisen.

1.3. **Growth, length, and age.** We have kept the 2013 formulation. Growth (only modelled for ages 8+) follows von Bertalanffy curves, with constant sex-specific k and t_0 for all adults of that sex, but individual-specific L_∞ that follows a Student's t_{12} distribution centred around a sex-specific mean (which is less sensitive than a Normal to outliers from measurement error). Growth is deterministic, given an individual L_∞. Formally, this is:

$$\mu_{sa} \triangleq \bar{L}_{\infty sa}\left(1 - \exp\left(-k_s\left(a - t_{0s}\right)\right)\right)$$
$$\mathbb{P}\left[\ell|sa\right] = F_{t12}\left(\left(\ell - \mu_{sa}\right)/\left(\mathrm{cvl}_s\mu_{sa}\right)\right) - F_{t12}\left(\left(\ell - \mu_{sa}\right)/\left(\mathrm{cvl}_s\mu_{sa}\right)\right)$$

where cvl_s is the CV of length-at-age and F_{t12} is the CDF of a standard Student's t_{12} distribution.

Since it is assumed that mortality rate depends on age rather than length, the population distribution (as opposed to sampling distribution) of length-at-age-and-sex keeps the same t_{12}-shape through the lifespan of each cohort, and individual fish maintain their "quantile" throughout adult life; in other words, a male SBT who was at the 15%le of length-at-age-8 will still be at the 15%le of length-at-age-28 twenty years later, if he survives. This considerably simplifies the calculation of HSP probabilities, described later.

Length classes are in 1cm intervals, with one grouped class at 150cm or below, and one grouped class at 200cm or above. For some calculations it is necessary to estimate contributions from animals below 150cm and above 200cm, so we also keep track of "mean conditional length" in those classes, e.g.

$$\mathbb{E}\left[L|L \leqslant 150\mathrm{cm}, s, a\right]$$

which is computed analytically from properties of Student's t-distribution, as per formulae in BGD2016.

Age is assumed to be measured accurately, both for juveniles (all assumed to be age 3, based on all samples coming from a specific range of lengths) and adults. Otolith age estimates are known to be

in error sometimes (usually by no more than 1 year), and it is also possible that some juvenile ages are wrong. In principle, ageing error can be accommodated in CKMR (and we have done so for other species), but it does greatly complicate the code. The effect of adult ageing errors on CKMR should not be that large; juvenile ageing errors would cause more problems, especially with mortality rate information from HSPs. The GT program should generate a lot more juvenile age-at-length data from tail vertebrae (a reliable indicator of age for immature SBT) so we will review that data when it becomes available.

Cohort-specific variations in growth have not been considered, though could be in principle.

1.4. **Fecundity and maturity.** We have kept the 2013 formulation; note that with the addition of HSPs, though, fecundity is no longer strictly coupled to selectivity (see below).

Males and females both reach maturity— i.e. to *potentially* breed successfully— at age 8. This is the youngest breeding age among the 76 POPs[2]. For animals 8yo and up, relative fecundity depends on length (in a sex-specific way) but not on age per se. Female fecundity (annual) is proportional to spawning-ground residence-time multiplied by daily fecundity; the latter (in terms of egg biomass released) has been estimated, as a function of length, from the histological analyses. Residence time (on some relative scale) is then estimated within the CKMR model, assuming a logistic relationship to length.

Male fecundity is again proportional to spawning-ground residence-time multiplied by daily fecundity, but we have no prior information on the latter. CKMR only gives information on annual total fecundity as a function of body size (by sex), so it is not possible statistically to separately estimate male residence time and male daily fecundity. However, we continue to use residence time as the *basis* for selectivity (perhaps modified by other factors), so retaining the distinction is somewhat useful. As for females, male residence time varies logistically with length. Male daily fecundity has been assumed independent of length in most options, since it appears that any such relationship is imprecisely estimated; presumably, the logistic residence-time/length relationship may already provide enough flexibility to describe male annual fecundity, especially since selectivity can now be adjusted separately. (The overall male fecundity/size relationship was also difficult to estimate in 2013, but the direct linkage to selectivity complicated

The 2013 analysis suggested that skip-spawning (in alternate years) might be the norm for younger SBT adults but not for older ones, and the newer data broadly supports that. The average time-interval between offspring birth and parental capture has increased— simply because the data series spans more years— so the direct interpretation of birth/capture gaps is more fraught; an adult that was "young" when it bred may no longer be "young" when recaptured. Analysing the phenomenon properly in this

[2]Actually, one parent in one post-2013 POPs does have an estimated age-at-breeding of 7. This is the only instance of an ostensible 7yo breeder in an enormous number of comparisons of 7yo adults (since each adult is involved in comparisons at all its ages between 2002 and the year before it was caught), while there are about six 8yo parents and far more at older ages; so, we suspect the 7yo estimate actually comes from otolith ageing error (estimated age 14 in this case). Even if this one instance is correct, it is abundantly clear that 7yo cannot make much of an overall contribution to SBT reproduction, so keeping age-at-maturity to 8 seems reasonable. Pushing down the age-at-maturity in the model introduces complications with length-frequencies that are just not worth the effort.

larger longer dataset would require a bespoke statistical model (albeit much simpler than, and completely separate to, the main CKMR analysis); this would be an interesting exercise.

Skip-spawning is an aspect of fecundity which is not specifically allowed for in the CKMR probabilities described below. In a reasonably long study, such as we certainly have now, the effects should cancel out; about half the probabilities are calculated too low and about half are calculated too high, so for individual fish the effects should broady cancel. At an aggregate level, the bulk of the reproductive output of SBT comes in any case from fish which are old enough to spawn every year.

While the CKMR framework of BSA2016** could certainly accommodate skip-spawning in principle, implementation would require much more complicated code. For SBT we expect there would be only minimal bias from using the current, much simpler, model.

2. POP equations

We have kept the 2013 formulation, which is a special case of the general ERRO ("Expected Relative Reproductive Output") approach explained in BSA2016. In words: the chance of any given adult being the parent of a specific juvenile, is the expected reproductive output of that adult in the year and place that the juvenile was born, divided by the expected TRO ("Total Reproductive Output") from all adults of the given sex at that year and place. The "expected" output needs to be calculated conditionally on all measured covariates for the adult; if the year and place of juvenile birth are not known but could take several values, then those values need to be integrated over conditional on what is known about the juvenile; and the art of CKMR is to express all that conditioning and integration correctly. For SBT, there is only one "place" (a single spawning ground) and juvenile age is (assumed) known, so all we need to worry about is the adult.

(1) Fecundity $\phi(s,\ell)$ is driven by length not age (provided age is 8 or more), but since the length-at-age distribution is constant over time, we can also compute an age-specific average by

$$\phi_{sa} \triangleq \sum_{\ell} \phi_{s\ell} \mathbb{P}[\ell|sa]$$

where $\mathbb{P}[\ell|sa]$ comes from the distribution of growth curves across individuals.

(2) Given the length, age, and sex of an individual adult in its year-of-capture[3], its length $\ell^*(s,\ell,a,y,y')$ in any previous year-of-interest y' can be back-calculated straightforwardly. The equations are too dull to include, but the steps are:
   (a) work out its individual L_∞ to match the observed length and age;
   (b) apply the individual growth curve at the age $a-(y-y')$ that the fish would have been in year y'.

For each combination of juvenile birth-year and certain adult covariates-at-capture (*not* including age, since age is only measured for a subset of genotyped adults), we compute probabilities based as follows. The ERRO calculation takes into account whether the adult was likely to be alive and mature at the

[3]There is not thought to be much length-growth during the spawning season, since adults lose weight on average, so date-of-capture within the season should not be important.

year of juvenile birth and, if so, its likely fecundity then:

$$\mathbb{P}\left[\text{POP}|bys\ell\right] = \sum_{a=8}^{A_{\text{MAX}}} \mathbb{P}\left[\text{POP}|bys\ell a\right] \times \mathbb{P}\left[a|ys\ell\left\{b\right\}\right]$$

$$\mathbb{P}\left[\text{POP}|bys\ell a\right] = \frac{\phi\left(s, \ell^*\left(s,\ell,a,y,y-b\right)\right) \times \mathbb{I}\left[y > b+3\right] \times \mathbb{I}\left[a-(y-b) \geqslant 8\right]}{\text{TRO}_{s,y-b}}$$

$$\mathbb{P}\left[a|ys\ell\right] = \frac{n_{sya}\mathbb{P}\left[\ell|sa\left\{y\right\}\right]}{\sum_{a'} n_{sya'}\mathbb{P}\left[\ell|sa'\left\{y\right\}\right]}$$

$$\text{TRO}_{sy'} = \sum_{a'} n_{sy'a'}\phi_{sa'}$$

When turning this into a log-likelihood, the first step is to group the comparisons by alike covariates $bys\ell$, so that instead of nearly 100,000,000 individual comparisons, we need only compute about 70,000 probabilities. Only a tiny proportion of pairwise comparisons actually yield a POP. The computationally-efficient way to form the log-likelihood is to first compute it as if there were no POPs at all, then loop over the POPs indivdually to adjust the log-likelihood for the facts that (i) this comparison between adult i and juvenile j did yield a POP after all, and (ii) the age of adult i was measured to be whatever it is (since all the adults involved in POPs are deliberately aged). That is:

$$\Lambda_{\text{POP}} = \sum_{bsy\ell} m_{\text{P}bys\ell} \log\left(1 - \mathbb{P}\left[\text{POP}|bys\ell\right]\right) +$$

$$\sum_{i,j:\ K_{ij}=\text{POP}} \text{logit}\mathbb{P}\left[\text{POP}\ |b_j y_i s_i \ell_i|\right] \log\left(\mathbb{P}\left[a_i|b_j y_i s_i \ell_i, K_{ij} = \text{POP}\right]\right)$$

$$\mathbb{P}\left[a|bys\ell, K = \text{POP}\right] = \frac{\mathbb{P}\left[\text{POP}|bys\ell a\right] \times \mathbb{P}\left[a|\left\{b\right\} ys\ell\right]}{\sum_{a'} \mathbb{P}\left[\text{POP}|bys\ell a'\right] \times \mathbb{P}\left[a'|\left\{b\right\} ys\ell\right]}$$

3. HSP equations

This follows the principles explained in **BSA2016. It is only necessary to compare *across* cohorts, not within them, so we can always distinguish the "first" and "second" juvenile in each comparison. Their joint HSP probability must be calculated by summing across all females (for MHSPs) or males (for PHSPs) alive at the time the first juvenile was born. Letting R_{ib} denote the *actual* (as opposed to expected or observed) reproductive output of animal i in year b, and supposing the two juveniles to be born in b_1 and b_2 respectively, the MHSP probability is simply

$$\mathbb{P}\left[K_{12} = \text{MHSP}|b_1 b_2\right] = \sum_{i \in \{\text{♀ alive at } b_1\}} \frac{R_{ib_1}}{\text{TRO}_{\text{♀}b_1}} \times \frac{R_{ib_2}}{\text{TRO}_{\text{♀}b_2}}$$

Note that the second output, R_{ib_2}, will be zero if the female i dies in-between b_1 and b_2.

Given a large adult population, we can replace the actual R's by their expected values for each female:

$$\mathbb{P}\left[K_{12} = \text{MHSP}|b_1 b_2\right] = \sum_{i \in \{\text{♀ alive at } b_1\}} \frac{\mathbb{E}\left[R_{ib_1} R_{ib_2}\right]}{\text{TRO}_{\text{♀}b_1}} \times \frac{1}{\text{TRO}_{\text{♀}b_2}} \tag{3.1}$$

Clearly, it is the *covariance* between an indidivual's reproductive outputs which matters, not merely the product of the two expected values. The key to success with HSPs in CKMR is to subsume as much as possible of that covariance into specific covariates of adults (length, age, etc), so that the bulk of the "covariance" is encompassed into a sum of products of expected values, which themselves are computed as in the POP calculations. Any remaining unaccounted sources of covariance will evidently cause "bias" in the number of HSPs found. Almost all plausible mechanisms, e.g. infertility, would lead to positive covariance, so that HSPs will occur *more* frequently than a naive calculation would indicate, and a corresponding abundance estimate would be biased downwards.

Note that "purely random" variability in individual R in any single year, e.g. through "sweepstake reproduction" and "lucky litters", does not matter; it is only *systematic* variability that needs to be accommodated. In the case of SBT, we can easily allow for adult length and growth by adapting the POP formulae; it would take a major mis-specification in our assumptions about growth trajectories (which are, basically, that bigger-than-average fish stay that way) to lead to much unaccounted covariance. Widespread infertility also seems a priori unlikely. Perhaps the most likely source of unaccounted covariance would be if some adults tend persistently to breed offspring who avoid the GAB in summer. To allow for such infelicities whatever their cause, the model can incorporate estimable scaling parameters α_{MHSP} and α_{PHSP} which act as multipliers on abundance when calculating HSP probabilities. If the model works as we hope, i.e. capturing the important sources of persistent individual-level fecundity, then these α_{HSP} should be close to 1. They might be substantially less than 1 if there are phenomena we have overlooked, so that there are more HSPs than "expected". But they cannot reasonably be much greater than 1.

With SBT, the HSP formulae do need to account for systematic variation due to (i) adult age; (ii) individual growth curves within age, in that if You are at the 15%le of length-at-age in year b_1 then You will also be at the 15%le in year b_2, and (iii) death. Point (ii) can be handled efficiently by summing across a fixed number Q of evenly-spaced quantiles of the t_{12}-distribution of length-at-age, so that $\{\ell\,(q\text{♀}a) :\ q \in \{1 \cdots Q\}\}$ approximates an equiprobable set of lengths-at-age. The other terms are already available from the POP calculations. The overall probability for an MHSP becomes:

$$\mathbb{P}\left[\text{MHSP}|b_1 b_2\right] = \frac{1}{\alpha_{\text{MHSP}}} \times$$
$$\sum_{a=8}^{A_{\max}} \left\{ \frac{n_{♀b_1 a} \phi_{♀a}}{\text{TRO}_{♀b_1}} \times \right.$$
$$\left. \frac{1}{Q} \sum_{q=1}^{Q} \left\{ \frac{\phi\left(\ell\left(q♀a\right)\right)}{\phi_{♀a}} \times \phi\left(\ell\left(q♀, a + (b_2 - b_1)\right)\right) \right\} \times \right.$$
$$\left. \exp\left(- \sum_{y=b_1}^{b_2 - 1} z_{♀ya+(y-b_1)} \right) \right\} \times$$
$$\frac{1}{\text{TRO}_{♀b_2}}$$

The extension to PHSPs is trivial, and we can also define $\mathbb{P}\left[\text{HSP}\right] \triangleq \mathbb{P}\left[\text{MHSP}\right] + \mathbb{P}\left[\text{PHSP}\right]$ since the two types of HSP are mutually exclusive; full-sibs across cohorts will be inconceivably rare in a random-mating population with millions of adults.

SBT have high haplotypic diversity in mtDNA, so much so that two HSPs which share an mtDNA haplotype are almost certainly an MHSP rather than a PHSP; if they have different haplotypes, of course, they must be a PHSP (see **main text). This lets us treat the mtDNA evidence as definitive about MHSP vs PHSP, and simplifies the log-likelihood calculation quite a lot; for all other species that we have looked at, it has been both necessary and tedious to take into account the observed haplotypes for each HSP when computing the log-likelihood.

The final step for HSPs is to note that the measured kinship in HSP comparisons is not necessarily the true HSP status, but rather the fact of whether the PLOD (see **main text) is above or below a certain threshold. That threshold is chosen on purely genetic grounds (i.e. before fitting the CKMR model) to eliminate any serious risk of false-positive "HSPs" from pairs that are unrelated or more weakly related. However— and in fact consequently— it is quite possible by chance that a true HSP will fall below the threshold and become a false-negative. This is accommodated by allow for a pre-estimated false-negative rate, in this case about 10%, so that the demographic HSP probabilities calculated above are all adjusted downwards to allow for false-negatives (i.e. reduced to about 90% of their nominal value) before computing the HSP log-likelihood, which is based on observed numbers of *definite* HSPs.

After all that, each pairwise comparison between juveniles is a "trinomial" event with outcomes MHSP/PHSP/UP, and the comparisons can be aggregated across birth-years to yield an overall HSP log-likelihood as a sum of multinomials:

$$\begin{aligned}\Lambda_{\mathrm{HSP}} = \sum_{b_1} \sum_{b_2 > b_1} \{ & m_{\mathrm{H}b_1b_2} \log\left(1 - \mathbb{P}\left[\mathrm{HSP}|b_1b_2\right]\right) + \\ & \#\mathrm{MHSP}_{b_1b_2} \log \mathbb{P}\left[\mathrm{MHSP}|b_1b_2\right] + \\ & \#\mathrm{PHSP}_{b_1b_2} \log \mathbb{P}\left[\mathrm{PHSP}|b_1b_2\right]\}\end{aligned}$$

4. Length/sex frequency data, and selectivity

This is the change from the 2013 model; , and it required more effort than actually incorporating the HSPs. As noted in the main text, selectivity in the Indonesian fishery appears to have varied substantially in some years between 2002 and 2014, with clear bumps and dips that cannot be explained merely by "overdispersion" at the level of independent 1cm length classes— the approach used in 2013, where the overdispersion was estimated from a "model-free" pre-analysis. Ignoring these selectivity shifts, or treating them just as overdispersion, would tend to overweight the LSfreq data in the overall log-likelihood, potentially compromising the fit to the other datasets (POPs, HSPs, A@LS).

The new model incorporates instead a year-specific random-effect spline which . The recorded LSfreq data $n^{\mathrm{LS}}_{\ell sy}$ is currently represented as Poisson count data with:

$$\log \mathbb{E}\left[n^{\mathrm{LS}}_{\ell sy}\right] = \mathrm{offset}\left(\log n_{sy\ell}\right) + \beta^{\mathrm{LS}}_y + \mathrm{tresid}_{s\ell} + \sigma_0\left(s, \ell\right) + \sigma_y\left(\ell\right) \tag{4.1}$$

$$n_{sy\ell} = \sum_{a=8}^{A_{\max}} n_{sya} \mathbb{P}\left[\ell|sa\right]$$

where β^{LS} is an intercept, $n_{sy\ell}$ and $\mathbb{P}\left[\ell|sa\right]$ come from the population dynamics, σ_0 is an overall selectivity following a prescribed functional form with estimable fixed-effect coefficients, and $\sigma_y\left(\ell\right)$ is the annual random-effect curves. The latter are chosen to be Duchon splines (ref**Wood 2016 Generalized Additive Models edition 2), which can be set up to penalize first-derivative penalties so that their default "preference", in the absence of any data, is to be constant across length; better-known choices such as cubic splines will tolerate any linear trend without penalty, which allows too much freedom. The term $\sigma_0\left(s, \ell\right)$ can be specified in any reasonable way; options considered so far are constant (leading to asymptotic selectivity driven by residence time only, as in the 2013 model), and continuous-piecewise-linear with kinks at 165cm and 190cm. The latter allows dome-shaped selectivity, as described in the main text.

The ADT software currently only permits one random-effect variance to be estimated internally inside a model, and that has been used up here for recruitment variability. Consequently (and as with LSfreq overdispersion in 2013), the spline variability has to be pre-estimated. This was done by comparing LSfreqs (aggregated across sexes) in adjacent years; population-dynamics-driven changes should be slow enough that the main difference across a single year would be due to selectivity shifts. To avoid having to specify any model for underlying length-frequency in the preliminary analysis, we fit a Gaussian GAM with Duchon splines to differences between successive annual *proportion*-at-length (i.e. normalized

by total LSfreq sample size in each year), with approximate weights computed by the Delta-method to account for Poisson variability. The spline variance from this model should transfer directly to equation (4.1). The preliminary analysis is fitted twice, starting first from an odd-numbered year and then from an even-numbered year, and the two estimated spline variances are averaged.

Because true variability in population length frequency is ignored, this type of preliminary analysis will tend to *slightly* underweight LSfreq data in the main model (although, as noted in the main text, the LSfreq data actually still seem to be getting too much weight, and the model for sex frequency is clearly not adequate yet— there is some more work to do). Neverthless, for a species like SBT with fairly slow turnover, it should be a simple and general way to allow for "uninteresting but unignorable" annual shifts in selectivity.

5. Age-at-length-and-sex data

Since selectivity is assumed to depend on length but not on age, each age-at-length-and-sex datum can be treated as an independent size-1 multinomial variable, with probabilities given by

$$\mathbb{P}\left[a|\ell s\right] = \frac{n_{sya} \times \mathbb{P}\left[\ell|sa\right]}{\sum_{a'} n_{sya'} \times \mathbb{P}\left[\ell|sa'\right]}$$