Marine Biodiversity Hub

National Environmental Science Programme

A national assessment of the status of White Sharks.

Barry Bruce[1], Russell Bradford[1], Mark Bravington[2], Pierre Feutry[1], Peter Grewe[1], Rasanthi Gunasekera[1], David Harasti[3], Rich Hillary[1], and Toby Patterson[1]

[1] CSIRO Oceans and Atmosphere; [2] CSIRO & Data 61; [3] NSW Department of Primary Industries

February 2018

CSIRO

Cover Image: Copyright Andrew Fox 2009 www.rodneyfox.com.au

Enquiries should be addressed to:

Barry Bruce and Rich Hillary
CSIRO Oceans & Atmosphere. GPO Box 1538, Hobart, Tasmania, 7001

Project Leader's Distribution List

The Hon Josh Frydenberg	Minister for the Environment and Energy
Finn Pratt	Secretary: Department of the Environment and Energy
Ashley Leedman/Lesley Gidding-Reeve Threatened Species Commissioner:	Department of the Environment and Energy: Marine and Freshwater Species Conservation Section: Wildlife, Heritage and Marine Division
Dr Sally Box	Department of the Environment and Energy - Threatened Species Commissioner
Amanda Richley	Parks Australia
Vic Peddemors	New South Wales Department of Primary Industries
Barry Hayden	South Australian Government
Carolyn Armstrong	Department of the Environment and Energy – ERIN
Natalie Moltschaniwskyj	New South Wales Department of Primary Industries

Preferred Citation

Bruce *et al*. 2018. A national assessment of the status of white sharks. National Environmental Science Programme, Marine Biodiversity Hub, CSIRO.

Copyright

Acknowledgment

This work was undertaken for the Marine Biodiversity Hub, a collaborative partnership supported through funding from the Australian Government's National Environmental Science Programme (NESP). NESP Marine Biodiversity Hub partners include the University of Tasmania; CSIRO, Geoscience Australia, Australian Institute of Marine Science, Museum Victoria, Charles Darwin University, the University of Western Australia, Integrated Marine Observing System, NSW Office of Environment and Heritage, NSW Department of Primary Industries. This work could not have been completed without the assistance of James Marthick (University of Tasmania, Menzies Institute for Medical Research).

Important Disclaimer

The NESP Marine Biodiversity Hub advises that the information contained in this publication comprises general statements based on scientific research. The reader is advised and needs to be aware that such information may be incomplete or unable to be used in any specific situation. No reliance or actions must therefore be made on that information without seeking prior expert professional, scientific and technical advice. To the extent permitted by law, the NESP Marine Biodiversity Hub (including its host organisation, employees, partners and consultants) excludes all liability to any person for any consequences, including but not limited to all losses, damages, costs, expenses and any other compensation, arising directly or indirectly from using this publication (in part or in whole) and any information or material contained in it.

Contents

EXECUTIVE SUMMARY

This report provides an overview of the key findings on Australasian white shark (*Carcharodon carcharias*) abundance and population dynamics which have superseded the initial estimates in Hillary *et al*. (2018). Both the submitted paper and the updated work detailed herein constitute the first robust estimates of Australasian white shark abundance and demographic rates (survival and trend) ever undertaken, and employ a variety of cutting-edge methods and novel data from across the Australian/New Zealand range of the species. Importantly, these estimates do not rely on the highly uncertain historical catch data for this species. To our knowledge, no similar study has been conducted worldwide. The results and methods employed, represent a step-change in capacity to assess otherwise difficult to monitor species, such as white sharks.

An update to the initial estimate was possible through access to a much larger dataset than was used for the previous estimate of total abundance for the eastern Australasian white shark population (NSW, TAS, VIC, QLD and New Zealand). This larger dataset, and the inclusion of mitochondrial DNA (which is inherited solely from the mother) permitted the exploration of both male and female adult shark population dynamics, and enabled the CSIRO to undertake a substantial update to the original estimates of both **adult** and **total** population size.

In addition, sufficient genetic samples had also accrued from the southern-western white shark population (SA, WA, western Victoria) to enable the first estimate of **adult** abundance for the southern-western population. Genetic and tagging data (both satellite and acoustic) have long suggested that these two populations are basically separate, with little reproductive cross-over. These updated analyses treat them as such.

Along with biological differences in the two populations, differences in sampling between eastern and southern-western populations also meant the available information was quite different. In the eastern Australasian population, sampling has been focussed on the juveniles; while for the southern-western population sampling has focussed on the sub-adult and young adult (male) life stages. This is largely due to the lack of a known southern juvenile aggregation area, such as exists for the eastern population along the mid-north NSW coast and also eastern Victoria, despite having undertaken surveys in regions where historically southern-western juvenile white sharks had been caught. Additionally, increased tagging effort and more extensive acoustic receiver coverage in the east has allowed for the estimation of juvenile survival (a key component when trying to estimate the total population size) for the eastern Australasian population, but not for the southern-western Australian population.

Somewhat counterintuitively, when looking at juveniles only (i.e. the eastern case) the key information source is actually related to adults via the genetic detection of closely related juveniles – specifically half-sibling pairs (HSPs), which share either a mother or a father. This information, collected from *juveniles* without ever having to sample the adults, tells us a great deal about the abundance and survival rates of *adults* over time. This is because, for two animals to be HSPs, their shared parent must have been alive for the interval between the two juveniles' respective birth years. For the southern-western population we look for siblings among the samples but also for what

are termed parent-offspring pairs (POPs) as these can also be highly informative on adult abundance in addition to the sibling pair data.

Eastern Australasian white shark population.

A total of 281 tissue samples with high quality DNA were available and were genotyped at thousands of places (termed loci) on the genome. This detailed genetic data is the key to finding the kin-matches (HSPs, POPs) used to estimate adult abundance. A stringent quality control process is used to ensure that spurious matches are not retained in the data set. A total of 214 samples and 3,097 loci passed the strict quality-control process, resulting in the identification of 73 half-sibling pairs and 23 full-sibling pairs.

Adult abundance for the eastern Australasian population was estimated to be N_{2017} = 750 (uncertainty range: 470 to 1,030). Survival (the chance of surviving from one year to the next) of adult males and females was not significantly different: 94% and 92%, respectively. Survival was thus assumed the same for both sexes and accurately estimated to be 93% – out of 100 fish known to be alive in one year, we would expect 93 of them to be alive the next year. The mitochondrial DNA revealed a 60:40 male-to-female sex ratio, consistent with no differences in survival between the sexes but where females mature later than males (as observed in the existing biological data).

The trend in abundance was not significantly different from zero (i.e. no trend and thus a stable population where births = deaths, on average). However, there was evidence for a slight decline over the 2000s (i.e. strong evidence against a high upward trend). Such a trend would be demographically consistent with the onset of protection (at the end of the 1990s) and the long time it would take for the effects of the various control programs and levels of fishing that existed pre-protection (which focused mostly on juveniles) to fully manifest themselves in the historical adult population. It is crucial to note that this work is able to inform on trends in adults *only*. Recent trends in the juvenile population (e.g. since protection their survival rates have possibly increased) are not expected to appear in the adult population until the future. We have addressed this point in the experimental design work.

From acoustic tagging data for juveniles in the eastern population we were able to estimate juvenile survival probabilities (73% with uncertainty range 63 to 83%). This estimate is plausible for a long-lived top predator, but it is based on data from only 49 sharks. As such, further refinement from larger datasets is advisable. Combining this information with adult abundance estimates enables an estimate of current (2017) **total** population size. The total population abundance estimate was N_{2017} = 5,460 (uncertainty range 2,909 – 12,802).

Southern-western Australian adult white shark population

A total of 271 tissue samples with high quality DNA from the southern-western population were available for analysis across thousands of loci on the genome. A total of 175 samples and 3,185 loci passed the strict QC process, resulting in the identification of 27 half-sibling pairs and 14 full-sibling pairs within the southern-western white shark dataset. At most 1 POP was found and these data were not used, given their limited information content.

The model estimated that adult survival was definitely in the 90% and above range, but the current data were not definitive in terms of a single robust estimate, so we explored survival estimates of 93

and 95% in the southern-western population. The mean estimate of **adult** abundance for the southern-western population was N_{2017} = 1,460 (uncertainty range 760 to 2,250). The precision of the southern-western estimate is lower than the east purely because there were less half-sibling pairs found (27 *vs.* 73), but it is clear that the southern-western adult population is estimated to be larger (by a factor of almost 2) than the eastern population. Despite conducting targeted surveys in South Australia and Western Australia, we were unable to locate any nursery grounds where juvenile white sharks could reliably be found and tagged. For this reason we have no juvenile survival estimates for this population, thus we cannot directly construct a total population abundance estimate, but if it is similar to juvenile survival in the east then it would be reasonable to assume that the southern-western white shark total abundance is larger than in the east. In terms of **adult** population trend, this was very similar to the east: not significantly different from zero, but tending slightly negative from the turn of the century.

Future sampling programs to detect population trends

This project has demonstrated that the close-kin mark-recapture approach can detect trends in **adult** abundance. However, given how long it takes animals to mature (12-13 years and above), recent possible trends in the juvenile population, specifically increases in survival and overall numbers post-protection of the species, will only really be detectable in the adult population in the future. To consider how further sampling will reduce uncertainty in trend estimates, we explored the effect of

1. numbers of annual sample sizes (numbers of fish to take genetic material from – 30, 65 and 100 *p.a.*)
2. sampling time-frame (5 and 10 years from 2017 onwards)

on the accuracy of any future trend estimates.

Based on options of 1 and 2, we calculated how probable these programs would be in terms of correctly identifying a trend in the adult population. Bounded by a likely maximum growth rate range of +/- 4%, we explored four possible growth rate scenarios: -2, 0 (no trend), 2 and 4% annual trends[1].

- For the eastern population a 5 year sampling program where 65 samples per year are collected would be able to robustly detect any of the explored trends in adult abundance.
- If the trends are more extreme (either higher or lower) they would be detected with increasing accuracy so we feel this level of sampling would be robust for the eastern population.

Close-kin mark-recapture uses the genetic information in the collected tissue sample to look at the parents of the animal from which the tissue sample came. In other words, it looks back in time. For example, the parents of all one year old sharks would be at least 12 years old, and therefore, the information content has its foundation at least 12 years in the past.

[1] To give some context to these numbers a 2% growth rate over 5 years is a 10% increase and over 10 years a 22% increase, a 4% trend over 5 years is a 22% increase and over 10 years a 50% increase.

For the southern-western population, where sampled animals are relatively older than in the eastern population, the information has its foundation even further back in time, closer to when protection was first implemented. If the typical age-range of southern-western individuals sampled in the future remains similar to those used in this study, it will take longer to accrue information on future adult abundance for the southern-western population than it has for the eastern population. So unless the problem of ready-access to juveniles in the southern-western population is addressed, a longer 10 year sampling program of at least 30 samples per year would be required to robustly detect any future trends in the southern-western white shark adult population.

1. INTRODUCTION

The white shark, *Carcharodon carcharias*, is a cosmopolitan species distributed throughout temperate oceanic, coastal and estuarine waters (Last and Stevens 2009; Harasti *et al*. 2017). In Australia it is most commonly found in southern waters ranging from southern Queensland through to Northwest Cape in Western Australia. Two genetically distinct populations have been identified in Australian waters (Blower *et al*. 2012): an eastern Australasian population encompassing the Australian eastern seaboard, New Zealand, and south-western Pacific Ocean including New Caledonian, Vanuatu and Tongan waters (Duffy *et al*. 2012); and a southern-western Australian population ranging from western Victoria to Northwest Cape, Western Australia and extending into Southern Ocean waters as far as 55 °South. Electronic tagging supports the existence of a two population model (Figure 1). With a few notable exceptions, white sharks are sparsely distributed throughout their range.

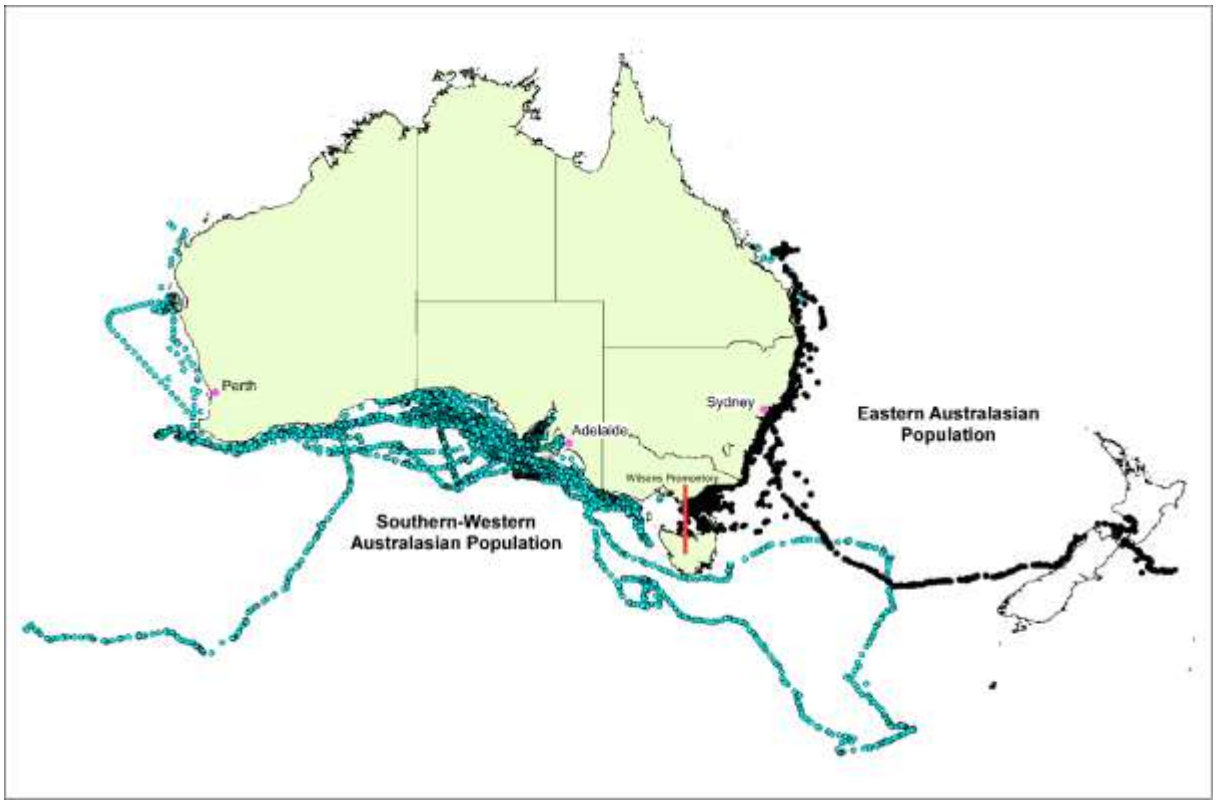

Figure 1: Support for the two Australian white shark population genetic model is provided by satellite tag deployments in Australia. Satellite tags deployed on the east coast of Australia are represented by filled circles; satellite tag deployments in South Australia and Western Australia are represented by solid lines. The approximate location of Wilsons Promontory is identified by the solid red vertical line.

White sharks display several classical K-selection traits (long-lived, high survival probability, and low fecundity) which make them sensitive to anthropogenic impacts. Perceived worldwide declines in population levels, largely due to fishing pressure, have resulted in protection, through various international and national legal instruments, being afforded throughout most of their range.

The white shark is listed as Vulnerable under the Australian Government's *Environment Protection and Biodiversity Conservation Act 1999* (EPBC Act) and is the subject of a national recovery plan (DSEWPaC 2013). A 2008 review of the 2002 Recovery Plan concluded that although progress had been made on a number of listed actions, there was still no effective way to estimate population size or trends and thus no effective way of determining if current recovery plan actions were having any beneficial effect. Recent public and political debate in New South Wales, Western Australia and South Australia, due to a series of shark attacks attributed to the species as well as claims of increasing interaction frequency, highlight the need to assess population status and trends in white sharks. This information is required to establish the efficacy of combined recovery actions, inform effective and defendable recovery and population rebuilding strategies, and provide a scientifically sound basis from which to inform policies that aim to balance conservation objectives and public safety.

This project has furthered the significant advances made under the National Environmental Research Program (NERP) by providing National population estimates for the species, advanced our knowledge on movement patterns and key areas of habitat use (hotspots) as well as developed strategies for the future monitoring of the species. The novel genetic and integrated modelling tools developed and trialled nationally on white sharks have links to other project initiatives under NERP and NESP. Combined they have, served as a testing ground for these new techniques to assess the condition and trend of Australia's threatened species populations for which conventional data provide an inadequate base to do so. Importantly, combined with other NESP project initiatives on population status of marine species of conservation concern (e.g. project A1 Northern Australian hotspots for the recovery of threatened euryhaline species; project A6 Prioritisation of research and management needs for Australian elasmobranch species; project A9 A close-kin mark-recapture estimate of the population size and trend of the east coast Grey Nurse Shark), this project has contributed as a case study to identify how much information (and hence investment) is required to provide an adequate level of policy advice regarding the status of threatened marine species.

Refining the eastern Australasian total population abundance estimates, and applying the knowledge and techniques developed to achieve estimates of the southern-western adult population have been the focus of NESP research. This research established a national scale population assessment for white sharks in Australia. As part of this research, new Close-Kin Mark-Recapture (CKMR) tools were trialled (e.g. the identification of cousins) that have the potential to estimate population trend from existing genetic samples. This latter development, although currently uninformative, holds promise for more rapid and effective assessments historic population status of Australia's threatened species populations for which conventional data provide an inadequate base to do so.

Specifically, this project has provided a national assessment of the southern-western **adult** population abundance and an update of the total eastern Australasian population abundance and status in order to establish the efficacy of existing recovery actions and provide a scientifically sound and rational basis from which to inform policies that aim to balance conservation objectives and public safety.

2. APPROACH

Project A3 has used Close-Kin Mark-Recapture, a powerful combination of conventional mark-recapture theory and modern genetics (Bravington *et al*. 2016b) first successfully applied to southern Bluefin tuna, and subsequently applied to euryhaline elasmobranchs (NESP Project A1: Northern Australian hotspots for the recovery of threatened euryhaline species), and eastern grey nurse shark (NESP Project A9: A close-kin mark-recapture estimate of the population size and trend of east coast grey nurse shark). Specifically, genetic tags are used rather than physical marks or tags, and "recaptures" are of an animal's close relatives rather than of the animal itself. This allows for the use of samples from dead animals, and circumvents some problems of standard mark-recapture such as local site fidelity. DNA profiles of individuals, derived from tissue samples, are compared across all individuals in the samples using the most up-to-date genotyping techniques to find half-sibling pairs (HSPs). Given adequate sampling, the proportion of kin-pairs found, and their spread in space and time, can be used in a mathematically sound and transparent mark-recapture framework (Bravington *et al*. 2016a) to estimate adult abundance, movement patterns, and possibly trend.

3. MODEL DEVELOPMENT

3.1 Sampling

The sampling regimes for the eastern Australasian and southern-western white shark populations were different. Specifically, in the eastern population, sampling focussed on the juveniles at identified nursery grounds and those caught in the NSW and Queensland Shark Control programs. Existing tagging data (acoustic and satellite) indicated that the juveniles are well mixed throughout their geographic range, and that most juvenile sharks visited the nursery grounds. A more recent and more extensive dataset on movement and mixing in the eastern population is now available for the eastern population, but has not been analysed.

For the southern-western population, the sharks which could be sampled were generally sub-adult and young adult (male) sharks, primarily those aggregating at the Neptune Islands in South Australia and supplemented with samples from the Western Australian Department of Fisheries' tagging program. Again, electronic tagging data has indicated that sharks within the southern-western population freely mix throughout their range. Tagging data and genetic evidence support very limited mixing between the eastern and southern-western populations; hence the adult abundance in the two populations is modelled separately.

3.2 Eastern Australasian white shark population model.

A detailed description of the models is provided in the supplementary material (Appendix A).

The original population estimate provided in Hillary *et al.* (2018) combined the CKMR output for the adult demographic and population abundance estimate (using half-sibling pairs, HSPs) with juvenile survival rates to estimate total current population size. This report builds on the original model by incorporating additional samples and data from the mitochondrial DNA (mtDNA, which is maternally

inherited) to allow estimation of sex-specific parameters. The additional data spans a greater number of years (1984 – 2016) and a greater number of identified kin-matches. These data were initially analysed using the simple exponential growth model used previously; then in a more nuanced state-space model for abundance allowing the incorporation of sex-specific parameters to investigate time-varying trends.

3.3 Southern-western Australian white shark population model.

For the southern-western population the models used were the same as for those used in the east. However, the focus of the sampling on sub-adult and adult male sharks allowed for an initial screening for the presence of father-offspring pairs (FOPs). If present, these parent-offspring pairs (POPs) can be highly informative on adult abundance in addition to the FSP and HSP data. Greater detail on the specifics of the models used for the southern-western population are provided in Appendix A.

4. JUVENILE SURVIVAL ESTIMATION

Traditionally, survival rates are estimated using a mark-recapture experiment. However for species that are present at low densities and are wide ranging (e.g. white sharks) traditional mark-recapture experiments are uninformative due to the very low recapture rates. Acoustic tags which transmit a unique identification are one means of overcoming the low recapture rate if they are coupled with a network of acoustic receivers throughout the animal's distribution. The initial capture and tagging of an animal provides the "mark"; with acoustic detections providing the "recapture". Cessation of "recaptures" indicates mortality or emigration from the system. Due to the lack of samples from juvenile white from the southern-western population, estimating survival rates is only applicable to the eastern Australasian population.

Off the eastern seaboard of Australia there is a network of over 300 acoustic receivers (coordinated by Australia's Integrated Marine Observing System-IMOS) spanning the north-south distribution of the eastern white shark's range, and extending into New Zealand waters. This extensive network of acoustic receivers allowed shark movements to be monitored around the eastern and southern coasts of Australia. In 2007, the CSIRO began a white shark tagging program centred on the previously identified juvenile nursery ground off Port Stephens, NSW. In response to high tag shedding rates of external tags, from 2008 onwards all tags were placed internally, allowing detections through the life of the tag (6 – 8 years). To determine the representativeness of the acoustic tag data, the movements of a subset of sharks were monitored using fin-mounted satellite-linked radio tags (SLRT).

A total of 49 juvenile white sharks were fitted with internal long-life acoustic tags between 2008 and 2014. Of these sharks, 32 were also fitted with a fin-mounted SLRT. Over a seven year period there were 186,743 individual acoustic detections extending from Queensland to Tasmania and across the Tasman Sea to New Zealand, covering the full range of movements observed in the SLRT dataset. The satellite tracking data which is independent from the fixed acoustic listening station data, showed that all juveniles made extensive use of the nursery area along the Northern NSW coast. Noting the relatively small size of the tagging data set, this nonetheless suggests that the assumptions around mixing of juveniles throughout the population are reasonable.

The model used to estimate juvenile survival is a simple version of the Cormack-Jolly-Seber model (Seber 1982), where both juvenile survival and the observation probability are assumed to be time and age independent. Under these assumptions we estimated juvenile survival in the eastern Australasian white shark population to be = 0.73 (95% CI 0.62 – 0.81).

A detailed description of the model and derivation of juvenile survival rates is provided in Hillary *et al*. (2018).

The approach we have used here has clearly demonstrated how large scale acoustic arrays coupled with long-life acoustic transmitters can be employed to estimate survival for wide ranging animals present at low density levels. Some key uncertainties remain, however. In particular, our study used 49 animals, largely sourced from the 3 – 5 year age groups. The degree to which older (> 6+) age groups remain coastal, and thus observable within the coastal arrays, is unknown. If there are significant changes in distribution, there would be a need to incorporate age-based effects on observation probability. Individual random variation in shark movements and the uneven distribution of acoustic receivers also affect observation probability. The much larger dataset currently being collected by NSW Department of Primary Industries that includes older sharks would allow for a more sophisticated approach to the analysis and allow these and other uncertainties to be addressed.

5. POPULATION SIZE AND TREND

5.1 Eastern Australasian white shark population

Adult abundance was initially explored using the simple exponential growth model:

$$N_{s,t} = N_0 \, \zeta_s \exp(\lambda t),$$

where population growth rate (λ) was fixed at zero and the current sex ratio (ζ_s) was explicitly assumed to be 50:50. Using this base model, total **adult** abundance was estimated to be $N_{2017} = 690$ (CV = 0.18) and **adult** survival = 0.93 (SD = 0.02). By allowing population growth rate to be estimated, λ was estimated to be -0.03 with a SD = 0.03 (i.e. not statistically different from zero). However, **adult** abundance was now estimated to be $N_{2017} = 590$ (CV = 0.23). Using this simple model **adult** population growth rates are likely to have slightly declined or remained stable since the early-mid 2000s, and highly unlikely to be greater than 3% per annum.

Sequencing of mtDNA has allowed for the use of a sex-structured model incorporating separate parameters for male and female sharks (e.g. different sex ratios and multiple paternity). Using the state-space model, the estimate of population growth rate was the same as for the simple model and not statistically different from zero. **Adult** abundance using the sex-structured model was estimated to be $N_{2017} = 750$ (CV = 0.19) and **adult** survival = 0.93 (SD = 0.03). However, the sex ratio was 0.39:0.61 (F:M) with a CV = 0.16.

Knowing the total **adult** abundance, survival rate and sex ratio we can now examine total population abundance by incorporating additional age-specific demographic parameters (e.g. sex-specific age-at-50% maturity, juvenile sex ratio, and juvenile survival rates) into the revised Approximate

Bayesian Computation (ABC) algorithm detailed in Appendix A. The CKMR analysis estimated more adult males than females were present in the population. Male white sharks mature more quickly (reach maturity at a younger age) and transition to adult survival rates earlier than female white sharks, providing a clear pathway to having a different sex ratio amongst adults even if the juvenile sex ratio is 50:50. Based on this, we defined age-at-50% maturity as follows: males = *U*[8, 9, 10, 11], and females = *U*[12, 13, 14].

The **total population abundance** for the eastern Australasian white shark population was estimated to be 5,460 (2,909 – 12,802).

5.2 Southern-western Australian white shark population

The key to estimating total population abundance is an understanding of juvenile survival probabilities in the population under review. This is not known for the southern-western white shark population and we, therefore, focus on only estimating the abundance of the **adult** population.

The samples collected from the southern-western population consist of sub-adults and young adults (unlike the focus on juveniles in the eastern population). A focus on older (and male) white sharks opened up an additional potential data source in the southern-western population not available in the eastern dataset – the identification of father-offspring pairs (FOPs). The identification of parent-offspring pairs (POPs) can be highly informative in relation to adult abundance and, for non-lethal sampling regimes, potentially informative on adult survival rates (Bravington *et al*. 2016). Unfortunately the preliminary exploration of the southern-western dataset identified at most one FOP. At present this result is uninformative, however, with continued sampling more FOPs will be identified that could be used in future analyses.

In the first case when estimating adult abundance, survival and sex ratio, the population growth rate was fixed at zero. The southern-western **adult** population was estimated to be N_{2017} = 2,225 (CV = 0.22), survival = 0.999, and sex ratio = 0.53 (SD = 0.12). A survival rate of essentially 100% is clearly unrealistic and most likely stems from either a lack of data or a lack of contrast between comparisons over time. However, true survival is likely to be in the mid-low 90% as was observed for the eastern population. Using an adult survival rate fixed at 0.93, the **total adult** population was estimated to be N_{2017} = 1,330 (CV = 0.22); and when adult survival rate was fixed at 0.95, **total adult** population size was estimated to be N_{2017} = 1,580 (CV = 0.22). In both cases the sex ratio did not differ from the original estimates.

The mean estimate of **adult abundance**, using a freely estimated survival rate, for the southern-western population was N_{2017} = 1,460 (uncertainty range 760 to 2,250).

When investigating trend using the simple exponential growth model, the rate of **adult** population growth was again estimated to be -0.03, but with a larger SD (0.04), and not significantly different from zero. As with the eastern population this is suggestive of a weak downward trend from the 2000s onwards, and is not consistent with a strong upward trend in the population.

6. STRATEGIES TO GUIDE MONITORING

6.1 Detecting population trends

The CKMR approach is capable of detecting trends in **adult** abundance. Detecting trend in the juvenile component requires a longer sampling period due to the time taken (12 – 13 years) for these animals to mature and become detectable in the adult population. For the eastern Australasian population we estimate that a 5 year sampling program of around 65 tissue samples per year from juvenile white sharks would be sufficient to robustly detect trends in **adult** abundance of between -2 and +4% per annum.

For the southern-western Australian white shark population juvenile nursery grounds have yet to be identified; hence the focus of previous and current sampling effort is on larger (i.e. older) individuals. This in effect "back-dates" the information content in the new data and means more time is required to accrue information on trends in adult abundance. We estimate that a 10-year sampling program of approx. 30 tissue samples per year would be sufficient to robustly detect trends in **adult** abundance of between -2 and +4% per annum for the southern-western population.

6.2 Improving estimates of juvenile survival

Juvenile survival rates are an integral part of estimating **total** abundance and projecting likely future trends in the adult component of the population. The dataset used for estimating juvenile survival for the eastern population was based on a small number of individuals and a restricted juvenile age range. Since 2015 the NSW DPI have launched an expanded tagging program, using both acoustic tags and SLRTs, with the number of tagged sharks approaching 300 (in 2017). Analysing this expanded dataset would greatly improve estimates of **juvenile** survival for the eastern population as well as potentially inform the key uncertainties around observation probability.

As outlined above, we had very few tissue samples from juvenile sharks for the southern-western population. This limits our ability to apply integrated population models to estimate total abundance for this population. Effort aimed at identifying pupping grounds and juvenile nurseries west of Wilsons Promontory should continue. A number of approaches are possible: 1. Pop-off satellite tagging of mature female sharks and juvenile sharks at the Neptune Islands in South Australia; 2. Greater focus on surveying areas identified as potential nurseries; 3. continued tagging operations in South Australia and Western Australia.

6.3 Spatial and Habitat Analysis

Despite the results of this study indicating stable or possibly declining abundance in both the eastern Australasian and southern-western Australian white shark populations, shark attack rates in Australia have risen over recent years. A greater understanding of the fine-, medium-, and broad-scale white shark distribution and movement patterns, including links to environmental predictors, may allow us to start to understand the reasons behind increasing human-shark interactions and to develop predictive models of shark relative abundance and fine-scale distribution. Even if there are predictable relationships between shark distribution and environment, discerning these from

background noise and individual variability in shark behaviour will require larger data sets. The recently collected NSW DPI tagging dataset is much larger than the tracking data employed in this study. If analysed in detail it is likely to provide a sound basis to determine the presence of environmental drivers and inform on patterns and variation in distribution and movement for the eastern Australasian white shark population in general. The available data for the southern-western population is currently insufficient for this purpose.

7. PEER-REVIEWED PUBLICATIONS

Bruce, B and Bradford R. (2015). Segregation or aggregation? Sex-specific patterns in the seasonal occurrence of white sharks *Carcharodon carcharias* at the Neptune Islands, South Australia. *Journal of Fish Biology* **87**: 1355-1370. doi: 10.1111/jfb.12827.

Abstract: The seasonal patterns of occurrence of male and female white sharks *Carcharodon carcharias* at the Neptune Islands in South Australia were reviewed. Analyses of a 14 year data series indicate that females seasonally aggregate in late autumn and winter coinciding with the maximum in-water availability of lactating female long-nose fur seals and seal pups. During this period, observed male:female sex ratios were similar; whereas during late spring and summer, males continued to visit, but females were rarely recorded. There was no evidence for segregation by sex or size at the Neptunes, but the highly focused seasonal pattern of occurrence of females compared with the year-round records of males suggests that there are likely to be differences between the sexes in overall distribution and movement patterns across southern Australia. It is suggested that foraging strategies and prey selection differ between sexes in *C. carcharias* across the life-history stages represented and that sex-specific foraging strategies may play an important role in structuring movement patterns and the sex ratios observed at such aggregation sites. Differences between sexes in distribution, movement patterns and foraging strategies are likely to have implications for modelling the consequences of fisheries by-catch between regions or jurisdictions and other spatially or temporally discrete anthropogenic effects on *C. carcharias* populations. Such differences urge for caution when estimating the size of *C. carcharias* populations based on observations at pinniped colonies due to the likelihood of sex-specific differences in movements and patterns of residency. These differences also suggest a need to account for sex-specific movement patterns and distribution in population and movement models as well as under conservation actions.

Burgess, G. H., Bruce, B. D., Cailliet, G. M., Goldman, K. J., Grubbs, R. D., Lowe, C. G., MacNeil, M. A., Mollet, H. F., Weng, K. C. and O'Sullivan, J. B. (2014) A Re-Evaluation of the Size of the White Shark (*Carcharodon carcharias*) Population off California, USA. *PLoS ONE* **9**(6): e98078. https://doi.org/10.1371/journal.pone.0098078.

Abstract: White sharks are highly migratory and segregate by sex, age and size. Unlike marine mammals, they neither surface to breathe nor frequent haul-out sites, hindering generation of abundance data required to estimate population size. A recent tag-recapture study used photographic identifications of white sharks at two aggregation sites to estimate abundance in ''central California'' at 219 mature and sub-adult individuals. They concluded this represented approximately one-half of the total abundance of mature and sub-adult sharks in the entire eastern North Pacific Ocean (ENP). This low estimate generated great concern within the conservation community, prompting petitions for governmental endangered species designations. We critically examine that study and find violations of model assumptions that, when considered in total, lead to population underestimates. We also use a Bayesian mixture model to demonstrate that the inclusion of transient sharks, characteristic of white shark aggregation sites,

would substantially increase abundance estimates for the adults and sub-adults in the surveyed sub-population. Using a dataset obtained from the same sampling locations and widely accepted demographic methodology, our analysis indicates a minimum all-life stages population size of .2000 individuals in the California subpopulation is required to account for the number and size range of individual sharks observed at the two sampled sites. Even accounting for methodological and conceptual biases, an extrapolation of these data to estimate the white shark population size throughout the ENP is inappropriate. The true ENP white shark population size is likely several-fold greater as both our study and the original published estimate exclude non-aggregating sharks and those that independently aggregate at other important ENP sites. Accurately estimating the central California and ENP white shark population size requires methodologies that account for biases introduced by sampling a limited number of sites and that account for all life history stages across the species' range of habitats.

Harasti, D., Lee, K. A., Liard, R., Bradford, R. and Bruce B. (2016). Use of stereo baited remote underwater video systems to estimate the presence and size of white sharks (*Carcharodon carcharias*). *Marine and Freshwater Research* **68**: 1391-1396. doi: 10.1071/MF16184.

Abstract: Stereo baited remote underwater video systems (stereo-BRUVs) are commonly used to assess fish assemblages and, more recently, to record the localised abundance and size of sharks. The present study investigated the occurrence and size of white sharks (Carcharodon carcharias) in the near-shore environment off Bennett's Beach, part of a known nursery area for the species in central New South Wales, Australia. Six stereo-BRUV units were deployed approximately fortnightly between August and December 2015 for periods of 5 h in depths of 7–14 m. Stereo-BRUVs successfully recorded 34 separate sightings of 22 individual white sharks. The highest number of individuals detected during a single day survey was eight. All C. carcharias observed on stereo-BRUVs were juveniles ranging in size from 1.50 to 2.46-m total length (mean +/- s.e., 1.91 +/- 0.05 m; n = 22). The time to first appearance ranged from 15 to 299 min (mean +/- s.e., 148 +/- 15 min). This study demonstrates that the use of stereo-BRUVs is a viable, non-destructive method to obtain estimates of the size and presence of white sharks, and may be useful in estimating relative abundance in near-shore environments where white sharks are known to frequent.

Harasti, D., Lee, K., Bruce, B., Gallen, C. and Bradford, R. (2017). Juvenile white sharks *Carcharodon carcharias* use estuarine environments in south-eastern Australia. *Marine Biology* **164**: 58. doi: 10.1007/s00227-017-3087-z.

Abstract: Estuarine environments are known to provide important feeding, breeding, resting and nursery areas for a range of shark species, including some which are considered dangerous to humans. Juvenile white sharks (<3 m) are known to frequent inshore environments, particularly ocean beaches, but their presence in and use of estuaries and coastal embayments is unclear. Given that estuarine environments are often surrounded by highly populated areas, understanding how white sharks use these environments will not only assist with their conservation management, but also inform public safety policies. The use of estuarine environments by acoustic-tagged white sharks was investigated from 2009 to 2015 at Port Stephens, New South Wales and Corner Inlet, Victoria, both of which adjoin known nursery areas for the species. Juvenile white sharks were detected within both estuaries, with 20 individuals recorded within the Port Stephens estuary, including four on one day. Only one tagged shark was detected within Corner Inlet; however, monitoring effort and local tagging in the area was more restricted. Detections in Port Stephens were predominantly from October to January and peaked in November. This study demonstrates that the footprint of known nursery areas for white sharks in eastern Australia should be expanded to include their adjacent estuarine environments. Consequently, there is clear potential for them to be exposed to a range of anthropogenic

estuarine impacts, and that human interactions are more likely over warmer periods (summer), when human use of such waterways is more prevalent.

Hillary, R. M., Bravington, M. V., Patterson, T. A., Grewe, P., Bradford, R., Feutry, P., Gunasekera, R., Peddemors, V., Werry, J., Francis, M. P., Duffy, C. A. J. and Bruce, B. D. (2018). Genetic relatedness reveals total population size of white sharks in eastern Australia and New Zealand. *Scientific Reports*.

Abstract: Conservation concerns exist for many sharks but robust estimates of abundance are often lacking. Improving population status is a performance measure for species under conservation or recovery plans, yet the lack of data permitting estimation of population size means the efficacy of management actions can be difficult to assess, and achieving the goal of removing species from conservation listing challenging. For potentially dangerous species, like the white shark, balancing conservation and public safety demands is politically and socially complex, often leading to vigorous debate about their population status. This increases the need for robust information to inform policy decisions. We developed a novel method for estimating the total abundance of white sharks in eastern Australia and New Zealand using the genetic-relatedness of juveniles and applying a close-kin mark-recapture framework and demographic model. Estimated numbers of adults are small (ca. 280-650), as is total population size (ca. 2,500-6,750). However, estimates of survival probability are high for adults (over 90%), and fairly high for juveniles (around 73%). This represents the first direct estimate of total white shark abundance and survival calculated from data across both the spatial and temporal life-history of the animal and provides a pathway to estimate population trend.

McAuley, R. B., Bruce, B. D., Keay, I. S., Mountford, S., Pinnell, T. and Whoriskey, F. G. (2017). Broad-scale coastal movements of white sharks off Western Australia described by passive acoustic telemetry data. *Marine and Freshwater Research* **68**: 1518-1531. doi: 10.1071/MF16222.

Abstract: Movements of 89 acoustically tagged subadult and adult white sharks (*Carcharodon carcharias*) were monitored off the south and west coasts of Western Australia (WA) between December 2008 and May 2016 by a network of up to 343 passive acoustic receivers. In all, 290 inter-regional movements, totalling 185 092 km were recorded for 73 of these sharks. Estimated rates of movement in excess of 3 km h^{-1} (mean 1.7 km h^{-1}; maximum 5.6 km h^{-1}) were common, even over distances of thousands of kilometres. Detections indicated that white sharks may be present off most of the south and lower west coasts of WA throughout the year, although they are more likely to be encountered during spring and early summer and are least likely to be present during late summer and autumn. There was limited evidence of predictable return behaviour, seasonal movement patterns or coordination of the direction and timing of individual shark's movements. Nevertheless, the data suggest that further analyses of movements in relation to ecological factors may be useful predictors of shark activity at local scales. It is hoped that these data may be useful for informing public safety initiatives aimed at mitigating the risks associated with human encounters with white sharks off the WA coast.

8. DATA ARCHIVE

Project data are being uploaded to the Australian Ocean Data Network data portal as they become available.
http://catalogue.aodn.org.au/geonetwork/srv/eng/metadata.show?uuid=8a41de4e-b649-45bd-b496-28250b627f77

All tissue samples collected as part of this project, or provided to CSIRO, have been processed to extract DNA. Tissue samples as well as extracted DNA have been archived at CSIRO Oceans & Atmosphere, Hobart.

9. ACKNOWLEDGMENTS

This project was funded by the Australian Government through the National Environmental Science Program's Marine Biodiversity Hub, with equal co-investment from the Commonwealth Scientific and Industrial Research Organisation. In NSW samples were collected under NSW DPI Animal Care and Ethics Committee permit number 12/07-CSIRO and NSW DPI Scientific Collection Permit P07/0099-6.0 (and their precursors). In South Australia samples were collected under SA Department of Environment, Water and Natural Resources Scientific collection permit U26255-4, Marine Parks Permit to Undertake Scientific Research MR00025-1, and Ministerial Exemption ME9902940 (including all precursors). An overarching Animal Ethics Permit was granted by the Tasmanian Department of Primary Industries, Parks, Water and Environment (AEC 22/2015-16), along with an authority to possess biological material from a listed species under the Living Marine Resources Management Act 1995 (permit 17109, and all precursors) and a Permit to Take Threatened Fauna for Scientific Purposes (permit TFA 17150, and all precursors).

This work would not have been possible without the immense assistance and collaboration shown throughout Australia and New Zealand. NSW DPI staff, including Dr Paul Butcher, Christopher Gallen, Roger Liard, Kate Lee, and many more, facilitated by Dr Natalie Moltschaniwskyj (Director of Fisheries Research, New South Wales Department of Primary Industries) have provided substantial support (people + infrastructure) in order to obtain samples and monitor acoustic arrays along the eastern Australian seaboard. Dr Malcolm Francis (National Institute of Water and Atmospheric Research, New Zealand) and Dr Clinton Duffy (Department of Conservation and University of Auckland) have provided tissue samples from white sharks tagged in New Zealand waters as well as provided acoustic data from tagged white sharks detected in New Zealand. Dr Rory McAuley, Silas Mountford, Ian Keay and Dani Waltrick (Department of Primary Industries and Regional Development, Western Australia) have provided samples from Western Australia and data from acoustic and satellite tagged sharks. Dr Paul Rogers (South Australia Research and Development Institute Aquatic Sciences) and Dr Charlie Huveneers (Flinders University, South Australia) have provided samples from South Australia and data from acoustic and satellite tagged sharks. Dr Jonathan Werry provided samples from Queensland. Finally (but definitely not least), we are immensely grateful to the South Australian Shark Cage Dive Industry, in particular Andrew Wright (Calypso Star Charters) and Andrew Fox (Rodney Fox Shark Expeditions) for their assistance in the field with collecting tissue samples and the maintenance of acoustic receivers at the Neptune Islands.

Professor Andre Punt (University of Washington, School of Aquatic Fisheries Science) and Dr Campbell Davies (CSIRO Oceans & Atmosphere) provided invaluable comment on the draft final report and technical appendix. Lastly we would like to acknowledge the staunch support and encouragement provided by Professor Nic Bax (Director Marine Biodiversity Hub National Environmental Science Programme) throughout the life of this project.

10. REFERENCES

Blower, D. C., Pandolfi, J. M., Bruce, B. D., Gomez-Cabrera, M. del C. and Ovenden, J. R. (2012). Population genetics of Australian white sharks reveals fine-scale spatial structure, transoceanic dispersal events and low effective population sizes. *Marine Ecology Progress Series* **455**: 229-244. doi: 10.3354/meps09659.

Bravington, M. V., Skaug, H. J. and Anderson E. C. (2016a). Close-kin mark-recapture. *Statistical Science* **31**, 259-275. doi: 10.1214/16-STS552.

Bravington, M. V., Grewe, P. and Davies, C. R. (2016b). Close-kin mark-recapture: estimating the abundance of Bluefin tuna from parent-offspring pairs. *Nature Communication* **7**, 13162; 10.1038/ncomms13162.

DSEWPaC (2013). Recovery Plan for the White Shark (*Carcharodon carcharias*). Available at http://www.environment.gov.au/biodiversity/threatened/recovery-plans/recovery-plan-white-shark-carcharodon-carcharias.

Duffy, C. A., Francis, M. P., Manning, M. J. and Bonfil, R. (2012). Regional population connectivity, oceanic habitat, and return migration revealed by satellite tagging of white sharks, *Carcharodon carcharias*, at New Zealand aggregation sites. In: *Global Perspectives on the Biology and Life History of the White Shark* (M. L. Domeier, Ed.) CRC Press, Boca Raton. pp. 301-318. ISBN: 978-1-4398-4840-1.

Last, P. R. and Stevens, J. D. (2009). Sharks and Rays of Australia, Second Edition. CSIRO Publishing, Canberra. ISBN: 978-0-674-03411-2.

Harasti, D., Lee, K., Bruce, B., Gallen, C. and Bradford, R. (2017). Juvenile white sharks *Carcharodon carcharias* use estuarine environments in south-eastern Australia. *Marine Biology* **164**: 58. doi: 10.1007/s00227-017-3087-z.

Seber, G. A. F. *The estimation of animal abundance and related parameters*. (Macmillan Publishing, New York, ed. 2, 1982).

APPENDIX A – TECHNICAL DETAILS FOR MARINE BIODIVERSITY HUB PROJECT A3: "*A NATIONAL ASSESSMENT OF THE STATUS OF WHITE SHARKS*".

Technical details for Marine Biodiversity Hub Project A3: *"A national assessment of the status of White Sharks"*

R.M. Hillary, CSIRO O & A T.A. Patterson, CSIRO O & A
M.V. Bravington, CSIRO Data61 R. Bradford, CSIRO O & A

February 7, 2018

1 Introduction

This report provides an updated estimate of the abundance and demographics of white shark (*Carcharodon carcharias*) in Australian waters. This builds upon previous estimates for the Eastern population using new data and provides the first estimate of abundance in the western population. In this document we detail the technical methodology used to calculate the estimates.

There are some differences in methods from the initial estimate for the east. Specifically, for the Eastern population, the model is essentially an augmentation of the original CKMR model (Hillary *et al.*, 2018), but now includes more data and also the availability of mitochondrial DNA (mtDNA, which is inherited only from mothers) allows estimation sex-specific parameters. This new information is used to update the whole-of-population estimate for the Eastern population initially explored in (Hillary *et al.*, 2018).

For the Western population, and due to the different nature of the sampling with respect to age, we explore a mixed approach which looks for both Parent-Offspring pairs (POPs - specifically focused on Father-Offspring pairs given the length distribution of the samples) and half-sibling pairs (HSPs). From this we provide an initial estimate of the adult western population (via the POP and HSP data).

Due to the lack of easy and reliable access to juveniles in the west we lack the requisite juvenile tagging data properly estimate juvenile survival. This is a key missing data set, preventing formal linkage of the juvenile and adult segments of the population - as was done in the East (Hillary *et al.*, 2018). Current evidence - genetic (Blower*et al.*, 2012), from both the archival and satellite tag data, and the initial joint analyses of Eastern and Western CKMR data (Hillary *et al.*, 2018), indicate that the Eastern and Western populations are effectively separate and with minimal reproductive cross-over, and we continue analysing them separately in this work. Generally speaking, there is obvious cross-over between the approaches to be applied in the East and the West. Nevertheless and for clarity, we present the results for each population in their own section - detailing the East first. Cross-references are made where useful.

2 Eastern population

We first outline the underlying adult population models and CKMR kinship probabilities, then how to use the mtDNA to estimate sex-specific parameters, the algorithm to estimate total population abundance, and then the results. While we go into data specifics later on, it is probably worth giving a high level overview of what data we have now, relative to when the first Eastern estimate was undertaken. In (Hillary *et al.*, 2018), after the genetic and fish quality control process, there were 75 fish (so 2,775 unique comparisons between fish) and around 2,200 SNP loci and 20 HSPs were identified. In the current data set we have 281 fish (although a number of these will be removed to due either being replicates/recaptures and via the detailed QC process) and around 15,000 initial SNP loci (of which many will also be removed but more than the original study). So we expect to have somewhere close to 3 times as many fish (meaning 8–9 times as many comparisons) and potentially more loci, both of which are expected to substantially increase the number of HSPs we are likely to find this time around.

2.1 CKMR Model structure

The CKMR (Bravington *et al.*, 2016) model used in this report is a modification of the original model (Hillary *et al.*, 2018), which is age-independent (in terms of both reproductive output, growth and survival) and covers the dynamics of the male, $N_{♂,t}$, and female, $N_{♀,t}$, abundance over time. The base case is a simple exponential growth model:

$$N_{s,t} = N_0 \zeta_s \exp(\lambda t), \tag{1}$$

where λ is the logarithmic rate-of-change and N_0 the initial abundance, ζ_s is the sex-ratio ($\zeta_{♀} = 1 - \zeta_{♂}$) and the annual adult survival probability is ϕ_s.

The updated CKMR data for the East actually covers a larger span of years (from 1984 to 2016) than was previously available, providing more kin-matches and hence, more information. This data supports investigation of a more nuanced state-space model for the abundance:

$$N_{t=0} = N_0, \tag{2}$$

$$\log N_t = \log N_{t-1} + \epsilon_t, \tag{3}$$

$$N_{s,t} = \zeta_s N_t, \tag{4}$$

$$\epsilon_t \sim N(0, \sigma_\epsilon^2), \tag{5}$$

where either ϵ_t (for a given value of σ_ϵ^2) or both ϵ_t and σ_ϵ^2 are process variances which can be estimated in a full random-effects framework (Skaug , 2002). This state-space formulation is a more complex model admitting more effective degrees of freedom via the term σ_ϵ^2. As such, this model can, in principle, deal with time-varying trends, which is not possible in the simple exponential form when estimating only λ.

2.2 Data likelihood for kin-matches

For the East, we conduct only juvenile-juvenile comparisons and look for full and/or half-sibling pairs. The key covariate for each $\{i, j\}$ juvenile comparison will be year and age-at-sampling: $\mathbf{z} = \{z_i, z_j\}$ and $z_\bullet = \{y_\bullet, a_\bullet\}$, but the key covariate is year of birth of cohort:

$c_\bullet = y_\bullet - a_\bullet$. For a more nuanced model where we do not have direct age estimates for each animal but we do have length and wish to integrate over the distribution of age given length the covariates are year and length-at-sampling. To define the various combinations of HSP probabilities let's take the within-cohort ($c_i = c_j$) case first. Taking the FSP case first:

$$\mathbb{P}\left(k_{ij} = FSP \,|\, \mathbf{z}, \eta\right) = \frac{\nu(1-\theta)}{N_{♀,c_i}},$$

and for MHSPs

$$\mathbb{P}\left(k_{ij} = MHSP \,|\, \mathbf{z}, \eta\right) = \frac{\pi_\eta^{\text{hsp}} \nu\theta}{N_{♀,c_i}}$$

and for PHSPs

$$\mathbb{P}\left(k_{ij} = PHSP \,|\, \mathbf{z}, \eta\right) = \frac{\pi_\eta^{\text{hsp}} \gamma}{N_{♂,c_i}}$$

Some specific details on the terminology used above:

- The ν term is the litter effect term - it parameterizes the potential for over-representation of within-cohort maternal kin (FSPs and MHSPs) given high heterogeneity in very early-life survival.
- The θ term is the probability that a mother will mate with more than one male in a given year i.e. multiple paternity litters.
- The term γ is similar to ν in scale (non-negative and can exceed one) and deals with the case where a father will mate with more than one female in a given year.
- The HSP false-neg probability π_η^{hsp} is needed to take care of the fact that we know we removed a certain fraction of true HSPs to avoid false positives in the final sample.
- The critical HSP PLOD η is the threshold where we expect to observe less than 1 non-HSP appearing above η; for this example $\pi_\eta^{\text{hsp}} = 0.94$ (with HTPs being the critical kin for defining where η was). For FSPs there was clear daylight in their PLODs vs non-FSPs so just assuming that $\pi^{\text{fsp}} = 1$.

We now deal with the cross-cohort case in which $c_i \neq c_j$; then for MHSPs (and assuming cross-cohort FSPs have effectively zero probability):

$$\mathbb{P}\left(k_{ij} = MHSP \,|\, \mathbf{z}, \eta\right) = \pi_{\eta}^{\text{hsp}} \frac{\phi_{\text{♀}}^{|c_i - c_j|}}{N_{\text{♀},c_j}}$$

.

The same general expression applies to cross-cohort PHSPs (just swap ♀ for ♂). . As shall be outlined below, there are no cross-cohort FSPs in the samples and we therefore explicitly fix the probability of these occurring to be zero in the models (both East and West). As a justification for this decision, consider if the probability of a cross-cohort HSP (male or female) is $\tilde{p}$, then the cross-cohort FSP will be approximately $\tilde{p}^2$. So the actual probability of something being above the η PLOD is likely to be close to $2\pi_{\eta}^{\text{hsp}}\tilde{p} + \tilde{p}^2$. If we omit the FSP part of this expression, we would be in effect under-estimating the probability of interest by a multiplier of

$$\left(1 + \frac{\tilde{p}}{2\pi^{\text{hsp}}}\right)^{-1},$$

Accordingly, if we assume that there are close to 350 males/females (700 adults) then, for an adult survival of 90% $\tilde{p}$ would be close to 0.001 so that multiplier will be about 0.9983. Contrast that with the inherent variance that we always ignore when invoking the classical survival equation: $N_{t+\tau} = N_t\phi^{\tau}$ (where for us ϕ is adult survival and τ the time between birth years of juveniles). That's the expected value of a binomial with sample size N_t and probability ϕ^{τ} so the CV on that would be

$$\sqrt{\frac{1-\phi^{\tau}}{N_t\phi^{\tau}}}$$

and again for $N_t \approx 350$, $\phi \approx 0.9$ and $\tau \approx 5$ that's a CV of around 4%. It is obviously small compared to the variance given the sample size of HSPs but a source of variance that looks a lot larger than the bias likely incurred from keeping things simpler in the model and assuming cross-cohort FSPs are treated as being zero probability events (and we didn't observe any).

2.3 Including the mtDNA

When it comes to the inclusion of the mtDNA, the general information source is: given the covariates, $\mathbf{z}$, and the critical PLOD, η, does an identified kin-pair share a haplotype or not? The approach we take is to split the cases into three groups:

- Case 1: Probability that you are above some PLOD value η', where $\eta' > \eta$, and are basically an FSP only (and therefore have to share a haplotype)
- Case 2: Probability that your PLOD is between η and η' (i.e. and HSP) and you share a haplotype ($h_2 = h_1$)
- Case 3: Probability that your PLOD is between η and η' (i.e. an HSP) and you do not share a haplotype ($h_2 \neq h_1$)

The are two steps to populating the probabilities that go along with these three cases: (i) calculate them 'raw' given the PLOD and covariates for each comparison above the critical PLOD, η; and (ii) normalise them to form a categorical distribution with the three cases as outlined above.

For case 1 (PLOD $> \eta'$) the raw probability is just $p_1 = \mathbb{P}(k = FSP \,|\, \mathbf{z}, \eta)$. For case 2 (PLOD $\in (\eta, \eta']$ and $h_2 = h_1$) is is defined as follows:

$$\begin{aligned} p_2 = &\mathbb{P}(k = MHSP \,|\, \mathbf{z}, \eta)\pi(h_2 = h_1 \,|\, h_1, k = MHSP) \\ &+ \mathbb{P}(k = PHSP \,|\, \mathbf{z}, \eta)\pi(h_2 = h_1 \,|\, h_1, k = PHSP), \end{aligned}$$

and for case 3 we have that

$$\begin{aligned} p_3 = &\mathbb{P}(k = MHSP \,|\, \mathbf{z}, \eta)\pi(h_2 \neq h_1 \,|\, h_1, k = MHSP) \\ &+ \mathbb{P}(k = PHSP \,|\, \mathbf{z}, \eta)\pi(h_2 \neq h_1 \,|\, h_1, k = PHSP). \end{aligned}$$

For MHSPs $\pi(h_2 = h_1 \,|\, h_1, k) = 1$ and $\pi(h_2 \neq h_1 \,|\, h_1, k) = 0$; for PHSPs $\pi(h_2 = h_1 \,|\, h_1, k) = \mathbb{P}(h_2)$ and $\pi(h_2 \neq h_1 \,|\, h_1, k) = 1 - \mathbb{P}(h_2)$. These three probabilities are then

normalised to sum to 1

$$p_m^{\text{mt}} = \frac{p_m}{\sum_{n=1}^{3} p_n}$$

and for every case, m, where we have an identified kin-pair above the critical PLOD η we add $\log\left(p_m^{\text{mt}}\right)$ to the log-likelihood, depending on which of three cases happens to cover that particular comparison that found the kin-pair.

The last bit of the likelihood (Bernoulli) will then be the nuclear DNA bit covering off on the probability of something being "above the η line":

$$\mathbb{P}\left(PLOD_{ij} > \eta \,|\, \mathbf{z}\right) = \sum_{k \in \mathcal{K}} \mathbb{P}\left(k \,|\, \mathbf{z}, \eta\right),$$

computed over all possible $\{i, j\}$ comparisons and where $\mathcal{K} = \{FSP, MHSP, PHSP\}$.

2.4 Estimation of total population abundance

In the original abundance paper (Hillary *et al.*, 2018) we employed an Approximate Bayesian Computation (ABC) algorithm to combine the CKMR adult demographic and abundance estimates with the juvenile survival estimates to estimate current total population size. The original algorithm was not sexually dimorphic, but the current CKMR models are, so it needs to be modified to be able to deal with more complex population model structures. We do, however, keep the general idea behind the original model: age-structured demography with an assumed stable age distribution and key statistics (specifically sex ratios at different times in the animals' life history) that are used to estimate approximate posterior distributions of total population abundance and demographic rates. Some useful variable terminology first:

- $N_{s,a}$: sex-specific abundance-at-age
- $\phi_{s,a}$: sex and age-specific annual survival probabilities
- λ: population growth rate
- $\beta_a = \bar{\beta} f_a$: age-specific birth rate (and $f_a \in [0, 1]$)

- $\ell_{s,a} = \prod_{i=0}^{a-1} \phi_{s,i}$: probability of survival to age a
- δ: average proportion of female pups at birth
- $\bar{R}$: total number of pups produced given female adult abundance

Classical population dynamics gives us the stable age distribution:

$$N_{♀,a} = \delta \bar{R} \ell_{♀,a} e^{-\lambda a},$$
$$N_{♂,a} = (1-\delta) \bar{R} \ell_{♂,a} e^{-\lambda a},$$

In terms of additional variables we need to fully describe the whole population-at-age we require:

- An ogive, $\chi_{s,a} \in [0,1]$, that defines the transition from the juvenile survival probability, ϕ^J (from the model fitted to the acoustic data), to the adult survival probability, ϕ^A (from the CKMR model), i.e.

$$\phi_{s,a} = \phi_s^J \chi_{s,a} + (1 - \chi_{s,a}) \phi_s^A$$

- A suitable model for $f_{s,a}$ which determines the age transition to adulthood (and can be different between the sexes) and, for the females, determines the age-specific birth-rate β_a along with the effective pupping rate $\bar{\beta}$.

For $f_{s,a}$ we employ a logistic model:

$$f_{s,a} = \left(1 + 19^{(a_s^{50} - a)/\tau^{95}}\right)^{-1},$$

where a_s^{50} is the age at which 50% "adulthood" (really maturity) is attained and $\tau^{95} = 3$ is the additional number of years after a_s^{50} at which 95% maturity has been achieved. Originally we used the maturity ogive (for females) as the transition vector $\chi_{s,a}$ but this

would not be appropriate for a genuinely sexually dimorphic model, so we employ a simple piecewise linear model for $\chi_{s,a}$:

$$\chi_{s,a} = \begin{cases} 1 & \text{if} \quad a < 6 \\ 1 - \frac{a-6}{a_s^{50}+\tau^{95}-6} & \text{if} \quad a \in [6, a_s^{50}+\tau^{95}] \\ 0 & \text{if} \quad a > a_s^{50}+\tau^{95}. \end{cases}$$

We chose age 6 as the point at which the age transition in survival occurs because this is currently the oldest age observed in the acoustic data and, given the lack of apparent age and sex-dependence in the juvenile survival estimates, this makes sense as the survival transition age at present.

From the CKMR we obtain estimates of not just ϕ_s^A and λ but total adult abundance, N^A, and the proportion of females ζ (and males $1-\zeta$). In terms of *total* adult abundance, we have that

$$N^A = \bar{R}\left(\sum_a e^{-\lambda a}\left(\delta f_{♀,a}\ell_{♀,a} + (1-\delta) f_{♂,a}\ell_{♂,a}\right)\right),$$

meaning $\bar{R}$ can be simply derived from all the other key abundance and demography parameters. It then follows that the effective pupping rate, $\bar{\beta}$, can be simply defined as follows:

$$\bar{\beta} = \frac{\bar{R}}{N_♀^A} = \left(\delta \sum_a f_{♀,a}\ell_{♀,a} e^{-\lambda a}\right)^{-1}$$

From the CKMR model we will have an estimate of the female-to-male sex ratio:

$$\widehat{\psi}^A = \frac{\zeta}{1-\zeta},$$

whereas that can be expressed in the age structured total abundance model as follows:

$$\psi^A = \frac{\delta}{1-\delta}\frac{\sum_a f_{♀,a}\ell_{♀,a}e^{-\lambda a}}{\sum_a f_{♂,a}\ell_{♂,a}e^{-\lambda a}}$$

Later on, a very useful ratio would be the sex ratio in some subset of the *immature*

population so assuming some suitable maximum age $\tilde{a}$ we have that

$$\psi^{\tilde{a}} = \frac{\delta}{1-\delta} \frac{\sum_{a=0}^{\tilde{a}} \ell_{♀,a} e^{-\lambda a}}{\sum_{a=0}^{\tilde{a}} \ell_{♂,a} e^{-\lambda a}},$$

In terms of the sex ratio in the juvenile population, in both the genetic samples (which are all juvenile and mostly less than 6–7 years of age) and the acoustic data, it is essentially 50/50 so our "estimate" would be $\widehat{\psi^{\tilde{a}}} = 1$.

The parameter vector in the modified ABC algorithm would be $\boldsymbol{\theta} = \{N^A, \phi^A, \phi^J, a_s^{50}, \lambda, \delta\}$. The prior distribution would be constructed as follows:

1. The key CKMR parameters (N^A, ϕ^A, and λ if estimable) are drawn from their asymptotic posterior distribution as obtained from the CKMR model
2. Juvenile survival ϕ^J prior defined as lognormal and parameterised using the log-scale mean ($\log(0.73)$) and CV (0.07) obtained from the acoustic data-driven mark-recapture survival model
3. The maturity parameters are assigned a uniform prior over a discrete set of possible integer values for a_s^{50}
4. The female pup fraction δ is assigned a beta prior, $B(a, b)$, where $a = b = 400$ so that the prior distribution of δ has a 95%CI of 0.46–0.54

The simple ABC rejection algorithm follows Marjoram *et al.* (2003) in the construction of an approximate posterior distribution for $\boldsymbol{\theta}$. Approximate Bayesian Computation, or ABC, is a powerful inference method when strict likelihoods cannot be defined. Classical likelihood driven inference posits a likelihood function for the model parameters, $\boldsymbol{\theta}$, given the data, D: $\ell(D \mid \boldsymbol{\theta})$. With a prior distribution for the parameters, $p(\boldsymbol{\theta})$, we can then define the *posterior* distribution of the parameters conditional on the data via Bayes' theorem:

$$\mathbb{P}(\boldsymbol{\theta} \mid D) = \frac{\ell(D \mid \boldsymbol{\theta}) p(\boldsymbol{\theta})}{\ell(D)},$$

where $\ell(D) = \int \ell(D \,|\, \boldsymbol{\theta}) p(\boldsymbol{\theta}) \mathrm{d}\boldsymbol{\theta}$. The ABC approach comes into its own when we cannot really construct (or calculate) $\ell(D \,|\, \boldsymbol{\theta})$, and in our case it is the construction issue we face. In the ABC paradigm, we relax the "likelihood given the data" requirement to one where we can simulate a particular (and hopefully informative) statistic D' from the model (defined by $\boldsymbol{\theta}$) and define a suitable distance metric $\rho(D, D')$. For a given tolerance level, $\varepsilon > 0$, if $\rho(D, D') \leq \varepsilon$ then we adjudge the given parameter vector $\boldsymbol{\theta}$ as acceptable, and not if otherwise. From this idea, we can define a very simple accept/reject algorithm to draw samples of $\boldsymbol{\theta}$ from an *approximation* to the posterior distribution $\mathbb{P}(\boldsymbol{\theta} \,|\, D)$.

The algorithm proceeds as follows:

1. Generate a proposal $\boldsymbol{\theta}^*$ from the prior $p(\boldsymbol{\theta})$
2. Simulate D'^* from the model defined by $\boldsymbol{\theta}^*$
3. Calculate the distance $\rho(D, D'^*)$
4. If $\rho(D, D'^*) \leq \varepsilon$ accept $\boldsymbol{\theta}^*$; if not, return to first step

By repeating the above algorithm until we have obtained a suitable number, we have generated our sample of $\boldsymbol{\theta}$ from the approximate posterior of interest. For our modified example we use two key pieces of information: the juvenile, $\psi^{\tilde{a}}$, and adult, ψ^A, sex ratios and each one is assigned a specific tolerance ($\varepsilon^{\tilde{a}}$ and ε^A, respectively). In the original ABC algorithm only the growth rate information was used, given it was a sexually combined model with an implicit 50/50 male/female sex ratio assumed (Hillary *et al.*, 2018).

One parameter (the proportion of adult females, ζ) is not estimated in the ABC algorithm, but is the key piece of information in one of the ABC algorithm's descriptive statistics, ψ^A. If ζ was strongly correlated to N^A, ϕ^J, or λ this might be problematic as we are using the CKMR estimate (and the variability thereof) to set both ψ^A and ε^A. However, perhaps not surprisingly, total adult abundance, survival and population growth show little to no correlation with ζ in the CKMR estimates. Also, though N^A plays no direct role in the ABC accept/reject algorithm it is absolutely vital in the estimation of $\bar{R}$ and, hence, total population abundance. Naturally its correlation with ϕ and λ is preserved given that the proposal/prior distribution for these three parameters is their asymptotic posterior distribution estimated in the CKMR model.

2.5 Exploratory data analyses

For the Eastern population we had 281 fish and over 14,000 SNP loci. As done originally (Hillary *et al.*, 2018) we follow a strict QC process to obtain both fish and loci that are consistent with established population genetics assumptions (such as Hardy-Weinberg equilibrium):

1. An initial sweep through the thousands of loci to get initial estimates of the minor (MAF) and null (NAF) allele frequencies and discard uninformative loci (with very low or high MAFs). Also to, where feasible, remove SNPs that are most likely on the same fragment of DNA (and therefore co-inherited and breaking the assumption of independent loci)

2. Detailed checks on the statistical properties of the remaining informative loci are done and loci that fail this process are discarded from future analyses

3. Checks are then done on the fish themselves (specifically their heterozygous properties) and fish which fail these checks are discarded from future analyses (and allele frequencies re-estimated post fish removal)

4. Repeat fish (either repeat samples or genuine subsequent "recaptures") are removed and the allele frequencies re-estimated

5. Full-sibling pairs (FSPs) are located

6. Half-sibling pairs are located, while taking care not to do comparisons between fish that are known to be in an FSP

7. Using the allele frequency model a strict false positive cut-off value for the statistic (pseudo log odds or PLOD) that differentiates an HSP from an unrelated pair (UP) to ensure there is less than 1 expected false-positive HSP (that could be a half-aunt/uncle of half-cousin). Once this cut-off is set calculate the resultant false-negative probability given the false-positive cut-off PLOD to ensure that we do not introduce bias into the abundance estimates when we exclude a certain percentage of true HSPs to avoid lesser-kin pair contamination

8. With our final set of fish and identified kin pairs retain only fish that have both a year and length or age-at-sampling (the key covariates for the CKMR model)

At the end of this process we had 214 fish, 3097 loci, 73 HSPs (and a false-negative probability of 0.94 so we are expecting to be losing around 6% of our true HSPs), and 23 FSPs (all the pairs born in the same cohort as far as we can tell and two triplet groups i.e. 3 litter mates and all the FSPs had the same haplotype which they must do). The mean gap between birth years for the UPs is 7.4 and for the HSPs it is 5.5 which suggests that survival is, similar to the original estimates, high for adults (which is also backed up by the identification of an HSP with a 22 year birth year gap). Figure 1 gives a more detailed graphically summary of the time between birth years of the 73 HSPs used in the CKMR analyses. In terms of mtDNA in the HSPs 36% of the HSPs *do* share a haplotype while 64% *don't* suggestive of a sex ratio bias towards males in the adults.

Figure 2 details a summary of HSP PLOD statistics as used to identify the half-sibling pairs. The top picture details the histogram of the UP plod on the left of the figure (with its predicted mean the blue dotted) line and the far smaller bump of HSPs on the left of the figure (and their predicted mean in magenta). For the UPs the predicted mean was within 1 standard error of the empirical mean, and the bottom of Figure 2 details how the predictive distribution of the UP PLOD (the red line) matches the empirical histogram. In particular whether the left-sided (to the left of the mean) distributions match closely, which they do. The empirical variance was slightly (*ca.* 12%) higher than the predicted variance, but the main point is that the allele frequency model is capturing the distribution fairly well and can be used to estimate the key false-positive cut-offs for the HSPs as outlined in Hillary *et al.* (2018).

2.6 Results

In terms of the simple exponential growth model outlined in 1 where no mtDNA was included and the sex ratio was explicitly assumed to be 50/50 current and λ fixed at zero total adult abundance, N^A_{2017} was estimated to be 690 with a CV of 0.18 and ϕ^A estimated at 0.93 (s.d. 0.02). When trend was estimated it was estimated to be $\lambda = -0.03$ with an s.d. of 0.03, the ϕ^A estimates were exactly the same and $N_{2017} = 590$ with a CV of 0.23. The trend

estimate, while slightly negative, was not estimated to be statistically significantly different from zero (likelihood ratio test) and, for the simple trend model, at best we can say that the current trend in the *adult* abundance is highly unlikely to be greater than 3%.

When including the mtDNA the estimate of trend was both the same and still statistically insignificantly different to zero so we omit any further estimates of models assuming an estimable exponential trend. When now estimating male and female parameters (including the female proportion, ζ, multiple paternity, $\boldsymbol{\theta}$, and multiple male partner, γ, parameters) the estimate of current total abundance is $N_{2017} = 750$ with a CV of 0.19, the ϕ^A estimates were exactly the same, and $\zeta = 0.39$ with a CV of 0.16. When permitting totally independent free parameters for male and female adult survival the female estimates were 0.92 (s.d. 0.03) and the male estimates were 0.94 (s.d. 0.04) so no clear evidence of different survival between the sexes.

When exploring a nuanced trend structure as defined in the state-space model (SSM) in 2 when estimating both the annual random effects, ϵ_t, and their variance σ^2_ϵ the variance was estimated to be very small $\sigma^2_\epsilon < 1e-6$ and the random effects ϵ_t were, while mostly negative, very very close to zero. All other parameter estimates coincided near exactly with their $\lambda = 0$ counterparts as one would expect. When permitting the model some freedom to move ($\sigma_\epsilon = 0.05$) there was a small but clear downward trend in total adult abundance over the last decade and a half (from about 2000 onwards) but - as with the exponential trend - it was clearly not statistically significant. So, after these various exploratory model analyses, the general summary for current Eastern white shark is:

- Current total adult abundance is around 700–750 sharks with a CV of just under 20% (versus around 500 sharks with a CV of almost 35% as detailed in Hillary *et al.* (2018))
- Survival is consistent across the sexes and well estimated at 0.93 (s.d. 0.03) (versus 0.98 with a s.d. of 0.05 in the original CKMR model)
- Sex ratio is currently around 60/40 in terms of male/female
- Adult abundance trend is not significantly different from zero, but with indications of a slight negative trend over the 2000s - strong evidence against any kind of current

high increasing trend in adult abundance though

To estimate total population size we need to fully define all the relevant variables and settings for the revised ABC algorithm as outlined in Section 2.4. The only prior distributions that require definition are those for the age-at-50% maturity (a_s^{50}) and these are defined as follows: $a_{\male}^{50} = U[8, 9, 10, 11]$, and $a_{\female}^{50} = U[12, 13, 14]$. The choice of these priors was motivated by the estimated difference in sex ratio between males and females (with more males estimated). Females are known to reach 50% maturity at around 12 to 13 years old and males are also thought to mature earlier. This is one clear pathway to having a different sex ratio even while males and females have essentially the same survival rates: even with a 50/50 juvenile sex ratio males both mature and transition faster from juvenile to adult survival rates. The tolerance for the adult female to male ratio was chosen to be $\epsilon^A = 0.05$ given the estimate and variability of ζ, and the tolerance for the juvenile female to male ratio up to and including age 6 was $\epsilon^{\tilde{a}} = 0.025$ given it is based on hundreds of observed animals and very likely more accurate than the adult sex ratio estimates.

The ABC algorithm was run to obtain 1,000 samples from the approximate posterior for $\boldsymbol{\theta} = \{N^A, \phi^A, \phi^J, a_{\female}^{50}, a_{\male}^{50}, \delta\}$ with λ fixed at zero given the CKMR model analyses. The key process variable of interest is total population abundance, $N_t^{\text{tot}} = \sum_s \sum_a N_{s,a,t}$, and we assume $t = 2017$ for reporting purposes. The median and 80% CI for N_{2017}^{tot} was 5,460 (2,909–12,802), the effective pupping rate β had a median and 80%CI of 4.3 (1.5–12.9), the mean number of 0-group $\bar{R}$ had a median and 80%CI of 1,200 (442–3,890), Figure 3 shows the posterior summaries for $a_{\female}^{50}$ and $a_{\male}^{50}$, and Figure 4 summarises the sex-specific survival probabilities-at-age.

3 Western population

For the Western population, the results in general will be more preliminary relative to the Eastern population because (a) this is the first attempt at estimating population size in the West, and (b) as we shall see there are less data to be analysed. The major focus and deliverable is on estimating adult population abundance and survival, given we currently have no direct estimates of juvenile survival probabilities which are the key to the estimation

of total population abundance. We do, however, explore some potential scenarios (assuming some level of comparability between survival rates of juveniles in the East and West) and what that might mean for total population abundance in the West.

3.1 CKMR Model structure

The population models explored (1 & 2) are the same, as are the probabilities for finding FSPs and HSPs but there is a further potential data source that *could* be present in the Western data: father-offspring pairs (FOPs). There is a small (around 10%) of the samples that are potentially old enough to be parents themselves, unlike the juveniles that make up the Eastern samples. Additionally, at least more recently, an increasing number of the samples are alive after the genetic samples have been taken further increasing the chances of finding FOPs in the samples. Generally speaking parent-offspring pairs (POPs) can be highly informative in relation to adult abundance and, for non-lethal sampling, potentially informative on adult survival also (Bravington *et al.*, 2016).

The FOP probability given a comparison between a male adult i and juvenile j is relatively simple to formulate, but we need to add an additional variable to the male adult covariate vector $z_i = \{y_i, a_i, k_i\}$, where y_i and a_i are the same as before but k_i denotes the life-status at sampling where $k_i = 1$ for alive and $k_i = 0$ for dead. The probaiblity of the $\{i, j\}$ comparison being a FOP is defined to be:

$$\mathbb{P}\left(k_{ij} = FOP \,|\, \mathbf{z}\right) = \frac{\mathbb{I}(c_i + a_m < c_j)}{N_{♂,c_j}} \times \begin{cases} 1 & \text{if} \quad y_i > c_j \\ \phi_{♂}^{c_j - y_i} & \text{if} \quad y_i \leq c_j \quad \& \quad k_i = 1 \end{cases}$$

where a_m is the male age-at-maturity, which the father has to have obtained before the juvenile was born) hence the condition using the indicator function $\mathbb{I}()$.

3.2 Exploratory data analyses

The same rigorous QC process was undertaken for the Western data (with an intial 271 samples in the data and the same number of initial SNPs as for the East). After the removal of poor performing/uninformative loci, fish that failed the QC checks, and the removal of duplicate samples (of which there were proportionally more than in the East) we finished

with 175 fish and 3,185 loci. There were 27 HSPs above the false-positive critical cut-off PLOD (and a false-negative probability of 0.89 so we are expecting to be losing around 10% of our true HSPs). We found 14 FSPs (including one triplet litter-mate grouping) born within the same years and sharing a haplotype. The mean time between birth years for the UPs is 8.65 and for the HSPs it is 8.04 again suggestive of high survival rates for adults (backed up by one HSP separated by almost 30 years in terms of cohorts) - see Figure 5 for the HSP cohort separation summary for the West. In terms of summarising the mtDNA for the HSPs 11 of the 27 HSPs *did* share a haplotype and 16 *didn't* suggestive of if anything a slightly higher number of females relative to males - the opposite of the Eastern case, albeit more uncertain with notably fewer data points. In terms of FOPs there was at most 1 FOP in the samples, making it something to consider for the future, but not informative at the moment. One final issue is the potential to have Grandparent-Grandchildren pairs (GGPs) in the samples - they are genetically indistinguishable from HSPs and over long enough time frames and with high adult survival that could become an issue that needs accounting for. However, although some of the males sampled in the early years *could* - if non-lethally sampled - have potentially been grandparents of current juveniles, all of them were dead-at-sampling (and specifically the few older animals in the long-term HSP cohort differences) and so we can (for now) discount the likely presence of GGPs hiding in the current set of HSPs.

Figure 6 details a summary of HSP PLOD statistics as used to identify the half-sibling pairs. The top picture details the histogram of the UP plod on the left of the figure (with its predicted mean the blue dotted) line and the far smaller bump of HSPs on the left of the figure (and their predicted mean in magenta). For the UPs the predicted mean was within 1 standard error of the empirical mean, and the bottom of Figure 5 details how the predictive distribution of the UP PLOD (the red line) matches the empirical histogram. In particular whether the left-sided (to the left of the mean) distributions match closely, which they do. The empirical variance was slightly (*ca.* 10%) higher than the predicted variance, but the main point is that the allele frequency model is capturing the distribution fairly well and can be used to estimate the key false-positive cut-offs for the HSPs as outlined in Hillary *et al.* (2018).

3.3 Results

So, just using the FSP/HSP data to estimate the key adult demographic and abundance parameters, there is again a clear consistency between including and not including the mtDNA data, so we omit those case here. When estimating abundance, survival and sex ratio but keeping λ fixed at zero we estimated that $N^A_{2017} = 2,225$ (CV of 0.22), $\phi^A = 0.999$ (hitting upper boundary), and $\zeta = 0.53$ (s.d. 0.12). It is hard to believe that adult survival is essentially 100% and, as with the *Glyphis glyphis* CKMR models, having either not enough data, or not enough contrast over time between the comparisons, can give rise to these kinds of estimates even if true survival is in the mid-to-low 90% range (i.e. the same as the East). In terms of a profile likelihood estimate of the lower 95%ile of the distribution of ϕ^A this was estimated to be about 0.93 (same as the East) so we look at fixed values of both 0.93 and 0.95 for ϕ^A in the West. For $\phi^A = 0.93$ we estimate that $N_{2017} = 1,330$ (CV of 0.22); for $\phi^A = 0.95$, $N_{2017} = 1,580$ (CV of 0.22). In both cases ζ (or its s.d.) did not change from the original estimates. When estimating trend the point estimate, for the exponential model defined in 1 was also $\lambda = -0.03$, but with an standard deviation of 0.04 so even less significant than the Eastern estimate. When moving to the more detailed SSM defined in 2 a very similar pattern to the East emerges: estimates of σ^2_ϵ are *very* small and so no significant trend is estimated. When permitting some freedom to estimate the trend (fix $\sigma_\epsilon = 0.05$ as before) a weak downward trend is estimated from the 2000s onwards (around a 7% overall reduction in adult abundance over this time).

4 Future data collection scenarios

A key variable of interest of policy makers, managers and the public is the trend in the overall population. The CKMR analyses have been reasonably clear in relation to adult trend over the temporal coverage of the data: both Eastern and Western adult populations show hints of a downward trend since the early 2000s but not significantly so. This consistent trend fits well with the time-line of when protection of the species occurred (basically by the end of the 1990s) and the age-at-maturity for this species - in general it would take at least 10–12 years for juveniles to be entering the population (i.e. around 2010 onwards. If protection did what one would expect it to do - increase the survival probability of the immature fish

(these were the ones caught in various control programs prior to protection) - then those fish that were experiencing a better survival regime for most of their life would only very recently be entering the adult population. We have consistently estimated adult survival probabilities to be high (greater than 90%), so the main part of the adult population will be comprised over easily over 25–30 age classes at any one time. So, even if there was a marked increase in juvenile survival post-protection (and a subsequent increase in recent juvenile abundance) it will very likely take a number of years into the future for it to be statistically detectable in the adult-focused CKMR analyses. With the juvenile acoustic tagging data we have a well-established process for estimating juvenile survival, and with the CKMR models we now have a well-established process for estimating adult abundance, survival and trend. The question now is not really about how long will it take to estimate population trend in the adults - the analyses both East and West are now informative on this issue - but more about how long will it take to detect a trend in juvenile abundance *in the adult population* if such a trend exists.

This is not just a policy and management issue but also one about future data collection protocols and experimental design, in terms of the numbers of samples and a time-frame for their collection given the desired outcomes (i.e. estimation of trend). As part of the general developmental work on CKMR we have been developing specific methods that can be efficiently used to explore just these kind of experimental design analyses. Originally conceptually outlined in Bravington *et al.* (2016) it involves specifying the future experimental design parameters (in this case number of samples per year, number of years over which samples are taken, age distribution of the samples) and then using statistical theory to construct the uncertainty estimates in the key parameters of interest, given the experimental design. It is thus relatively simple to tailor the specific aspects of the experimental design to obtain target estimates of expected uncertainty (or relatedly of certainty in specific process variables such as the probability that the estimated trend is positive, for example).

The CKMR models we have explored for white sharks have a Bernoulli/binomial likelihood function as they relate to pairwise comparisons of animals and whether they are related in a specific way or not - e.g. half-sibling or parent-offspring pairs. The general idea is to construct the Hessian matrix for the key estimable parameters (which defines the approximate parameter covariance matrix) via an approximation to the Fisher information

(FI) at the base information level for each data source. For the HSP-based CKMR models explored herein that base level is the cohort/year of birth of each animal in the comparison. The Fisher information is defined in two ways. The first is in terms of the expected value of the square of the score (gradient of log-likelihood) at the maximum likelihood estimate:

$$\mathcal{I}(\boldsymbol{\theta}) = \int \left(\frac{\partial}{\partial \boldsymbol{\theta}} \ln \ell(X \mid \boldsymbol{\theta}) \right)^2 \ell(X \mid \boldsymbol{\theta}) \mathrm{d}X = \mathbb{E}^X \left[\left(\frac{\partial}{\partial \boldsymbol{\theta}} \ln \ell(X \mid \boldsymbol{\theta}) \right)^2 \right], \tag{6}$$

where $\ell(X \mid \widehat{\boldsymbol{\theta}})$ is the likelihood function. The Fisher information (see Appendix) can also be written in terms of the second derivative of the log-likelihood (the Hessian):

$$\mathcal{I}(\boldsymbol{\theta}) = -\mathbb{E}^X \left[\frac{\partial^2}{\partial \boldsymbol{\theta}^2} \ln \ell(X \mid \boldsymbol{\theta}) \right] \tag{7}$$

For a Bernoulli process with probability p, the FI is given by $1/(p(1-p))$ (Appendix); for a binomial process with sample size n this becomes $n/(p(1-p))$. For the Bernoulli process a conservative approximation to the FI is simply $1/p$ - for CKMR models where the probability of recapture is often very small this basically is a very minor under-estimation (and associated minor over-estimate of uncertainty) in the FI.

In our case, the probability of occurrence is actually a function of the parameters we are really interested in, $p(\boldsymbol{\theta})$. The Fisher information of the parameter vector (see Appendix) is defined as follows:

$$\mathcal{I}(\boldsymbol{\theta}) = \mathcal{I}(p) \left(\frac{dp}{d\boldsymbol{\theta}} \right)^2 = \frac{1}{p} \left(\frac{dp}{d\boldsymbol{\theta}} \right)^2, \tag{8}$$

which, to improve numerical stability given small probabilities, can be shown (see Appendix) to be equivalent to the following:

$$\mathcal{I}(\boldsymbol{\theta}) = 4 \left(\frac{d\sqrt{p}}{d\boldsymbol{\theta}} \right)^\dagger \left(\frac{d\sqrt{p}}{d\boldsymbol{\theta}} \right) \tag{9}$$

This gives us the (approximate) Hessian matrix for the data at their base level covariates, $\mathbf{z}$. A very useful result is that the overall Hessian matrix, H, can be defined in terms of the sum over all the unique covariate groups of the individual Hessian matrices $H(\mathbf{z})$ as defined in (9):

$$H = \sum_{\mathbf{z} \in \mathcal{Z}} n^{\mathbf{z}} H(\mathbf{z}), \tag{10}$$

where $\mathcal{Z}$ is the set of possible covariates, and $n^{\mathbf{z}}$ is the number of observations of that covariate type in the data set. The parameter covariance matrix, $\Sigma_{\boldsymbol{\theta}}$, is the inverse of this overall Hessian matrix. Approximate estimates of the variance of parameters *derived* from $\boldsymbol{\theta}$ (such as total adult abundance over time) via some differentiable function $g(\boldsymbol{\theta})$ can be derived via the delta method:

$$\mathbb{V}(g(\boldsymbol{\theta})) \approx \left(\frac{dg}{d\boldsymbol{\theta}}\right)^{\dagger} \Sigma_{\boldsymbol{\theta}} \left(\frac{dg}{d\boldsymbol{\theta}}\right). \tag{11}$$

To make sure all these calculations are accurately done, we use automatic differentiation (AD) - specifically the `TMB` package in R - so that all gradients and derived quantities are calculated to machine precision.

4.1 CKMR model for design work

We have outlined the general theoretical framework we will use to derive all our relevant uncertainty estimates, given the design specifics, but we now outline the specifics of both the future population and experimental design model. Generally speaking:

- We focus on total adult abundance, N_t, adult survival ϕ^A, and future population growth λ
- We ignore within-cohort comparisons (where $c_i = c_j$) as these contain little to no actual information on the key parameters of interest, and also form a small fraction (*ca.* 5%) of historical comparisons in the data
- For the experimental design we assume that we have a fixed number of overall samples per year n^s, a number of years into the future for which we take samples, y^s, and an age distribution in the annual samples defined by π_a^s

For both the Eastern and Western CKMR models the recent trend in the adult population was not statistically different to zero, and we assume that this is true for the cohorts for which we have data (1983–2016). To account for future trends in adult abundance consider the following population model:

$$N_{t+1} = \begin{cases} N_t & \text{if} \quad t < \tilde{t}, \\ N_t e^{\lambda} & \text{if} \quad t \geq \tilde{t}, \end{cases}$$

where $\tilde{t}$ represents some future time at which the adult population trend begins. The cross-cohort probability for detecting an HSP (either maternal or paternal) is given by the following

$$\mathbb{P}\left(k_{i,j} = HSP \,|\, \mathbf{z}\right) = \frac{\left(\phi^A\right)^{|c_i - c_j|}}{\zeta(1-\zeta)N_{c_j}}, \tag{12}$$

where $\mathbf{z} = \{c_i, c_j\}$ and ζ is the proportion of females in the adult population (with $1 - \zeta$ the male counterpart). It is this probability in 12 that is used to construct the Hessian matrix specific to that covariate $\mathbf{z}$ as defined in 9, and we construct the overall parameter Hessian matrix by summing over all such unique covariate groups in the data, involving calculating the number of comparisons of that covariate type in the data, as defined in 10. To ensure that the future estimates are compatible with the past we integrate both the existing *observed* data (from which the underlying abundance and survival parameters were estimated), and *future* data (more specifically, the expectation thereof).

4.2 Future data collection scenarios

The actual samples are collected from both live and dead animals, so it is not possible to be totally definitive about how many samples can be collected each year, but there is clear information from field personnel on maximum and minimum numbers of annual samples that can be collected. A minimum level would be around 30 samples, with a maximum level of around 100 samples, but the indications are that the collection of 50–100 samples is clearly possible if the current staffing and infrastructure around the shark sampling and control programs in various states is maintained. In terms of the age structure in the samples, this was estimated from both the length distribution in the samples and the current length-at-age distribution as outlined in Hillary *et al.* (2018). Field personnel have outlined to us that it would be very difficult to focus the sampling towards specific age groups (i.e. towards younger or older animals outside of the mix that appear in the historical samples). Therefore, short of increasing the number of overall samples and then sub-sampling these

for the desired mix of ages, there is little we can really do to alter the age distribution of the samples, π_a^s.

Just to quickly explore whether the sub-sampling approach would make any real financial sense, the costs of genotyping a fish for CKMR purposes is around $30 a sample. If, as an example, 50% of the fish in the samples are aged 1–5 (with the other 50% 6–11) and we wanted to maximise the number of young fish to "front-end" our information in the data to the present and immediate future (older samples will yield information on historical adult dynamics) to try and better estimate current and future trend. Our maximum number of total samples is 100 so we could decide to collect all 100 and keep only the 50 that are in our required age-range. We have already invested the money to collect the maximum number of samples we can, so the only financial saving we would get is by not genotyping the 50 older animals (over a 10 year sampling program that would be a saving of $15,000) as the analyses costs exactly the same for 50 vs. 100 samples per year. So, we would be throwing away information by not genotyping the whole 100 samples per year and saving an amount that is likely a small fraction of the overall costs of the data collection and analyses in terms of staff and general infrastructure.

In terms of total annual samples we explore n^s values of 30, 65, and 100; for numbers of years to sampled for we explore y^s options of 5 and 10 years; the estimated prior age distribution in the Eastern and Western samples, π_a^s, is outlined in Figure 8. For the "true" trend in the future adult population we explore $\lambda \in \{-0.02, 0, 0.02, 0.04\}$ and assumed that the trend begins in the first future year of data so $\tilde{t} = 2017$. Given the three options for both n^s and y^s we have 6 cross-combinations and in terms of performance metrics we explore six options:

1. The approximate standard deviation of the estimated value $\widehat{\lambda}$
2. $\mathbb{P}(\widehat{\lambda} > 0)$: probability that the estimate of the true value of λ, $\widehat{\lambda}$, will be positive
3. $\mathbb{P}(\widehat{\lambda} < 0)$: probability that the estimate of the true value of λ, $\widehat{\lambda}$, will be negative
4. The approximate standard deviation of the estimated value of ϕ^A
5. The CV of the most recent (2017) estimate of total abundance

6. The CV of the the final year estimate of total abundance $(2017 + y^s)$

While simple these three statistics should give us a fairly robust indication of both how precisely we are estimating λ (the standard deviation) and the other key parameters, and how well we are estimating the qualitative nature of the true trend λ (the chances our estimate $\widehat{\lambda}$ will be positive/negative if the true trend λ is negative/positive).

n^s	y^s	λ	$\sigma(\widehat{\lambda})$	$\mathbb{P}(\widehat{\lambda} > 0)$	$\mathbb{P}(\widehat{\lambda} < 0)$	$\sigma(\widehat{\phi^A})$	CV (N_{2017})	CV (N_{fin})
30	5	-0.02	0.052	0.35	0.65	0.013	0.12	0.19
65	5	-0.02	0.032	0.27	0.73	0.01	0.08	0.11
100	5	-0.02	0.023	0.2	0.8	0.007	0.06	0.07
30	10	-0.02	0.016	0.11	0.89	0.008	0.09	0.12
65	10	-0.02	0.01	0.01	0.99	0.005	0.06	0.06
100	10	-0.02	0.007	0.01	0.99	0.004	0.04	0.04
30	5	0	0.053	0.5	0.5	0.013	0.12	0.2
65	5	0	0.033	0.5	0.5	0.009	0.08	0.11
100	5	0	0.024	0.5	0.5	0.007	0.06	0.08
30	10	0	0.017	0.5	0.5	0.009	0.09	0.13
65	10	0	0.01	0.5	0.5	0.006	0.06	0.07
100	10	0	0.007	0.5	0.5	0.005	0.04	0.05
30	5	0.02	0.055	0.64	0.36	0.013	0.12	0.2
65	5	0.02	0.034	0.72	0.28	0.01	0.08	0.11
100	5	0.02	0.025	0.79	0.21	0.007	0.06	0.08
30	10	0.02	0.018	0.87	0.13	0.009	0.09	0.14
65	10	0.02	0.011	0.97	0.03	0.006	0.06	0.07
100	10	0.02	0.008	0.99	0.01	0.005	0.04	0.05
30	5	0.04	0.056	0.76	0.24	0.014	0.12	0.21
65	5	0.04	0.034	0.88	0.12	0.01	0.08	0.12
100	5	0.04	0.025	0.94	0.06	0.008	0.06	0.08
30	10	0.04	0.019	0.98	0.02	0.01	0.1	0.15
65	10	0.04	0.011	0.99	0.01	0.007	0.06	0.08
100	10	0.04	0.008	0.99	0.01	0.005	0.04	0.05

Table 1: *Design summary statistics for the Eastern population.*

Table 4.2 has the detailed summary of all 6 key performance statistics, for all 24 of the options around the annual sample sizes and monitoring time-frame as well as the four "true" λ scenarios. We report the detailed probabilities for identifying if the trend was either positive or negative, but to assist in interpretation we can focus on the odds ratio (OR) for identifying the true trend. Two simple examples: if the true trend is positive and we we estimate the probability that the trend is positive to be 0.5, then or odds ratio is

n^s	y^s	λ	$\sigma(\widehat{\lambda})$	$\mathbb{P}(\widehat{\lambda} > 0)$	$\mathbb{P}(\widehat{\lambda} < 0)$	$\sigma(\widehat{\phi^A})$	CV (N_{2017})	CV (N_{fin})
30	5	-0.02	0.253	0.47	0.53	0.021	0.16	1.01
65	5	-0.02	0.109	0.43	0.57	0.014	0.1	0.43
100	5	-0.02	0.081	0.4	0.6	0.012	0.07	0.32
30	10	-0.02	0.038	0.3	0.7	0.014	0.12	0.32
65	10	-0.02	0.019	0.15	0.85	0.008	0.07	0.16
100	10	-0.02	0.013	0.07	0.93	0.007	0.05	0.11
30	5	0	0.258	0.5	0.5	0.021	0.16	1.03
65	5	0	0.111	0.5	0.5	0.014	0.1	0.44
100	5	0	0.082	0.5	0.5	0.011	0.07	0.33
30	10	0	0.039	0.5	0.5	0.013	0.12	0.33
65	10	0	0.02	0.5	0.5	0.009	0.07	0.16
100	10	0	0.014	0.5	0.5	0.007	0.05	0.12
30	5	0.02	0.263	0.53	0.47	0.021	0.16	1.05
65	5	0.02	0.114	0.57	0.43	0.014	0.1	0.45
100	5	0.02	0.084	0.59	0.41	0.011	0.07	0.33
30	10	0.02	0.041	0.69	0.31	0.014	0.12	0.35
65	10	0.02	0.02	0.84	0.16	0.009	0.07	0.17
100	10	0.02	0.014	0.92	0.08	0.007	0.05	0.12
30	5	0.04	0.268	0.56	0.44	0.021	0.16	1.07
65	5	0.04	0.116	0.63	0.37	0.014	0.1	0.46
100	5	0.04	0.086	0.68	0.32	0.01	0.07	0.34
30	10	0.04	0.042	0.83	0.17	0.014	0.12	0.36
65	10	0.04	0.021	0.97	0.03	0.009	0.07	0.18
100	10	0.04	0.015	0.99	0.01	0.007	0.05	0.13

Table 2: *Design summary statistics for the Western population.*

0.5/(1-0.5) which is 1; if our estimated probability is 0.67 then our odds ratio is 0.67/(1-0.67) which is 2.

If the true trend is negative (specifically -2% *pa.*) then for a 5 year future sampling program the approximate OR for identifying the correct nature of the trend are 2, 3 and 4 for 30, 65 and 100 annual samples, respectively. For a 10 year sampling program 30 samples per year outperforms 100 samples per year for 5 years; for 65 and 100 samples per year the true trend is fairly well estimated and we can be very sure (ORs of around 9 and above) if it is negative or not at the -2% per year level. For all future sampling regimes both ϕ^A and total adult abundance are well estimated, though better when the growth rate is well estimated - particularly total abundance in the final year of sampling. For zero future trend ($\lambda = 0$) and a 5 year sampling program the 65 and 100 annual samples will likely yield a

trend uncertainty envelope (ca. 95%) between ±6%. All other estimates (survival/current and future abundance) look acceptable and again improve with more sampling. For a small positive growth rate of 2% per year (a 10% increase in 5 years, 22% in 10) the 5 year sampling scenario yields similar results to the -2% trend scenario (ORs of around 2, 3 and 4 for the increasing annual sampling scenarios). The 10 year programs all do a better job (ORs close to 9 and above and basically definitive for 65 and 100 samples per year) at detecting a true but small positive trend. For a stronger positive trend of 4% per year (a 22% increase in 5 years, 50% in 10) and a 5 year program the 30, 65 and 100 samples per year do a fairly good job of probabilistically identifying the positive trend (approximate ORs of 3, 9 and 20, respectively). For a 10 year program the 30 samples per year option actually does a better job than the 65 samples over 5 years option in terms of identifying the positive nature of the trend (98% vs. 88%), at a decreased sampling cost, and with no appreciable trade-off in the estimates of both ϕ^A and total adult abundance (both now and in the future).

Table 4.2 details the key performance statistics for the Western population design analysis. A key difference with the East is that the apparent ages that can be sampled in the West is skewed towards older sub-adult animals (see Figure 8). A consequence of this is that, over a short (e.g. 5 year) future time-frame very little information relating to adult future dynamics will be accrued, which is not the case in the East. What is clear from Table 4.2 is that none of the 5 year sampling programs - no matter what the growth rate be it positive, negative or steady - would yield any really useful information on trend, for the reasons outlined in the previous sentence. Only one of the ORs approaches 2 (100 samples and a 4% positive trend) and most hover very close to 1. On the 10 year sampling time-scale, irrespective of the magnitude of the true trend, all the annual sample size scenarios yield ORs of at least 2 and often into the 4 and above range (even for 30 samples year scenario with a 4% positive trend).

5 Discussion & conclusions

The original (Hillary *et al.*, 2018) CKMR estimates of the Eastern adult population, and the whole-of-population estimates using the acoustic detection data to estimate juvenile

survival, have been with the larger number of available samples. For the first time an initial estimate of the Western adult population has been undertaken using a similar CKMR model as used for the East. Building on the updated Eastern estimates as well as the initial Western estimates, we have also undertaken a detailed design study which explores future data collection strategies (sample sizes and time-frames) which are logistically feasible and likely to be able to identify future trends in the adult population.

For the Eastern analyses after the detailed quality control process we had 214 fish genotyped a just over 3,000 well-performing and informative loci on the genome, thus making the robust detection of close-kin (specifically full and half-siblings) feasible. In all 73 half-siblings were found above the strict false-positive relatedness criterion, which will exclude lesser kin such as half-cousins and half-aunties and uncles appearing in the samples, and with year and age/length-at-sampling information. The resultant false-negative loss rate was 6% so the true number of half-siblings will be a little higher than 73 but in reality we lost very little in terms of true half-siblings. For the Western analyses we had 175 fish genotyped at almost 3,200 loci, and with year and length/age-at-sampling information. We detected 27 HSPs above the false-positive cut-off levels and the false-negative rate this time was 11% so, again, the true number of HSPs will be around 30. Given there are likely to be some adult males in the Western samples (they are clearly longer/older than in the Eastern samples) we also looked to see if there were any parent-offspring pairs that can also be used in CKMR models, but at most only 1 was found and, hence, we didn't explore this option further in this instance.

For the Eastern population, current (i.e. 2017) total adult abundance was estimated to be 750 individuals (CV 19%), with an accurately estimated annual survival probability (estimated to be the same for both sexes) of 0.93, and a male/female sex ratio of 60:40. Historical trend in the *adult* population was estimate to be slightly negative over the 2000s, but not significantly different to zero trend. A modification of the original model used to estimate total population size in Hillary *et al.* (2018) was developed to deal with males and females and potential sexual differences in the age-at-maturity, which are a clear potential driver for the observed 60:40 male/female sex ratio as males are thought to mature earlier than females. Current (i.e. 2017) *total* abundance was estimated to be around 5,500 animals, with a range between 3,000 and 12,800.

For the Western population current total adult abundance was estimated to be between 1,300 and 1,600 (both with CVs of 22%) for the two survival probability options of 0.93 and 0.95, respectively. The actual estimate of adult survival was very close to one, given more limited data than the East but clear long-term gaps (the largest being 28 years) between the birth years of the HSPs detected - a clear indication of high adult survival. The lower end of the likely estimates for the Western adult survival probability was around 0.93 so we explored this and 0.95 as two options. What is clear is that adult survival is very likely in the 90% and higher range. Sex ratio estimates for the West were actually very close to 50:50 for males and females. As with the Eastern results, current trend in the 2000s is very marginally negative but not significantly different to zero. We did not estimate *total* population size for the West, given the lack of enough juvenile tagging data to estimate sub-adult survival rates. However, if they are similar to the East then, given the adult population in the West is estimated to be larger than that in the East, one would expect that the total population in the West is also larger than the East. The only way to robustly assess this hypothesis would be to obtain Western juvenile survival estimates, and the only likely option for that is to tag more juveniles with acoustic devices and estimate survival as we did in the East.

Current estimates of *adult* population trend are estimated to be very slightly negative, but are not statistically distinguishable from zero. Given that protection for this species was fully implemented at the end of the 1990s, and assuming that protection increased juvenile survival rates (as they formed the majority of the animals being caught or killed), this result is not really surprising. The maturity-at-age for males (around 11–12) and females (13–16) means that the juveniles born after protection was put in place would really only be starting to enter the adult population in the last 5 years or so. Given very high adult survival rates (above 90%) there are also likely to be many, many age classes (between 25–30) that form the majority of the adult population. Given the likely age structure, the detectable demographic effect of these juveniles born post-protection, and even those born before then (and experiencing better survival rates since protection), will take a number of years to manifest in the adult population – which are the sole segments of the population about which CKMR is informative.

While we have estimates of *current* trend from our CKMR analyses, an obvious question

for policy is: how well can we estimate future population trend? From previous fieldwork we already have a very good idea for likely levels of minimum and maximum sample sizes that can be obtained per year, as well as the likely age-structure in those samples (for the East shorter/younger fish than in the West). Using this information we can perform the necessary detailed experimental design study to answer the question of how best to detect future trends in the adult population. In terms of annual sample sizes we looked at 30, 65 and 100 samples per year; considered sampling scenarios at 5 and 10 year time-frames beyond the present; for future trends we looked at -2%, 0%, 2% and 4% and focused on both how well we can estimate the true trend but also how well we can probabilistically categorise the trend - is it positive or negative?

In the East, if the true trend is small (e.g. $\pm 2\%$) then the 5 year sampling program scenarios are at least twice as likely to estimate the true nature of the trend (positive or negative) correctly as to get it wrong. For the 10 year scenarios this odds ratio of identifying the correct nature of the trend increases to a factor of almost 9 and above, for all the annual sample size scenarios. In terms of the survival and abundance estimates in all cases the precision of the future estimates improves on what has already been estimated in the current study. Therefore, it is mostly the issue of trend that would be the focus of discussions around future sampling. The key trade-off in the East will be time *vs.* cost - 30 samples per year for 10 years outperforms 100 samples per year for 5 years in this regard, and does so with 200 less samples. Does that saving in terms of genetics and annual sampling effort compare well with the cost of extending the time-frame of the sampling program for 5 more years? For the scenarios we explored, future sampling of at least 65 samples per year for at least 5 years is highly likely to deliver usable information on *future* adult trend (for the range explored) as well is maintain the accuracy of future abundance and survival estimates.

In the West the samples tend to be from older animals than in the East, and so the information content is "back-dated" more than in the East. So, even for 5 years of additional data collection at even 100 samples per year, we would not really accrue any strong information about the future adult population trend at all - most of the new samples were born prior to 2017 and so their information on adult dynamics (specifically trend) relates to this time-period. For the 10 year time-horizon all the annual sampling scenarios yielded odds ratios of just over 2 and higher, in relation to correctly identifying the nature of the true

trend. One issue for the lowest annual sampling level of 30 fish per year is that estimates of future abundance are likely to be much more uncertain than for the higher annual sample size scenarios. For the West, it appears that samples of at least 30 animals per year and for 10 years or so would likely be required both to say anything meaningful about future adult population trend, but to maintain the accuracy of future abundance and survival estimates this annual sample size would have to be around a 50 or more animals.

References

Blower, D. C., Gomez-Cabrera, M. C., Bruce, B. D., Pandolfi, J. M. & Ovenden, J. R. 2012. Population genetics of Australian white shark (*Carcharodon carcharias*) reveals a far more complicated breeding and dispersal biology than simple female-mediated philopatry. *Mar. Ecol. Prog. Ser.* **455**: 229–244.

Bravington, M. V., Skaug, H. J., & Anderson, E. C. 2016. Close-kin mark-recapture. *Stat. Sci.* **31**: 259–275.

Hillary, R. M. *et al.* 2018. Genetic relatedness reveals total population size of white sharks in eastern Australia and New Zealand. *Nat. Sci. Rep.* (submitted).

Marjoram, P., Molitor, J., Plagnol, V. & Tavare, S. 2003. Markov chain Monte Carlo without likelihoods. *Proc. Natl. Acad. Sci.* **100**: 15,324–15,328.

Skaug, H. J. 2002. Automatic differentiation to facilitate maximum likelihood estimation in non-linear random effect models. *J. Comput. Graph. Stat.* **11**: 458–470.

Appendix

Calculating the Fisher information via the second derivative option requires far more calculation than does the option via the square of the score (i.e. needing only first derivatives). The two methods defined in the main text are equivalent, and we show now why is this the case. Consider first the second derivative of the log-likelihood:

$$\frac{\partial^2}{\partial \boldsymbol{\theta}^2} \ln \ell(X \,|\, \boldsymbol{\theta}) = \frac{\frac{\partial^2}{\partial \boldsymbol{\theta}^2} \ell(X \,|\, \boldsymbol{\theta})}{\ell(X \,|\, \boldsymbol{\theta})} - \left(\frac{\frac{\partial}{\partial \boldsymbol{\theta}} \ell(X \,|\, \boldsymbol{\theta})}{\ell(X \,|\, \boldsymbol{\theta})} \right)^2 = \frac{\frac{\partial^2}{\partial \boldsymbol{\theta}^2} \ell(X \,|\, \boldsymbol{\theta})}{\ell(X \,|\, \boldsymbol{\theta})} - \left(\frac{\partial}{\partial \boldsymbol{\theta}} \ln \ell(X \,|\, \boldsymbol{\theta}) \right)^2$$

Taking the first part of the above equation and calculating its expectation over the data:

$$\mathbb{E}^X \left[\frac{\frac{\partial^2}{\partial \boldsymbol{\theta}^2} \ell(X \,|\, \boldsymbol{\theta})}{\ell(X \,|\, \boldsymbol{\theta})} \right] = \int \frac{\partial^2}{\partial \boldsymbol{\theta}^2} \ell(X \,|\, \boldsymbol{\theta}) \mathrm{d}X = \frac{\partial^2}{\partial \boldsymbol{\theta}^2} \int \ell(X \,|\, \boldsymbol{\theta}) \mathrm{d}X = 0.$$

The additive nature of the expectation, and the result in the previous equation, give rise the the equivalent definitions of the Fisher information:

$$\mathcal{I}(\boldsymbol{\theta}) = \mathbb{E}^X \left[\left(\frac{\partial}{\partial \boldsymbol{\theta}} \ln \ell(X \,|\, \boldsymbol{\theta}) \right)^2 \right] = -\mathbb{E}^X \left[\frac{\partial^2}{\partial \boldsymbol{\theta}^2} \ln \ell(X \,|\, \boldsymbol{\theta}) \right].$$

For the Bernoulli process, with probability p, the likelihood (and log-likelihood) of a binary outcome $X = \{0, 1\}$ is

$$\begin{aligned} \ell(X \,|\, p) &= p^X (1-p)^{X-1}, \\ \ln \ell(X \,|\, p) &= X \ln p + (1 - X) \ln(1-p), \end{aligned}$$

and the second derivative of the log-likelihood is given by

$$\frac{\partial^2}{\partial \boldsymbol{\theta}^2} \ln \ell(X \,|\, p) = \frac{X}{p^2} + \frac{1-X}{(1-p)^2}.$$

For the Bernoulli distribution we have that $\mathbb{E}(X) = p$, so the expected value of the second derivative of the log-likelihood (i.e. FI) is:

$$\mathcal{I}(p) = \frac{p}{p^2} + \frac{1-p}{(1-p)^2} = \frac{1}{p} + \frac{1}{1-p} = \frac{1}{p(1-p)} > \frac{1}{p},$$

which would make p^{-1} a conservative approximation to the Fisher information for Bernoulli processes (such as the kin detection probabilities in the CKMR models).

Here we detail exactly how we calculate the Fisher information matrix for the estimated parameters, given the original Fisher information is defined in terms of the underlying probability (which is a function of the estimated parameters). Assuming $p(\boldsymbol{\theta})$ is at least two-times differentiable in terms of $\boldsymbol{\theta}$ then, via the chain rule, we have that

$$\mathcal{I}(\boldsymbol{\theta}) = -\mathbb{E}^X\left[\frac{\partial^2}{\partial p^2}\ln \ell(X \,|\, p)\left(\frac{\partial p}{\partial \boldsymbol{\theta}}\right)^2 + \frac{\partial}{\partial p}\ln \ell(X \,|\, p)\frac{\partial^2 p}{\partial \boldsymbol{\theta}^2}\right].$$

Taking the second part of the expression in the above expectation we have that

$$\mathbb{E}^X\left[\frac{\partial}{\partial p}\ln \ell(X \,|\, p)\frac{\partial^2 p}{\partial \boldsymbol{\theta}^2}\right] = \frac{\partial^2 p}{\partial \boldsymbol{\theta}^2}\mathbb{E}^X\left[\ln \ell(X \,|\, p)\right] = \frac{\partial^2 p}{\partial \boldsymbol{\theta}^2}\frac{\partial}{\partial p}\left(\int \ell(X \,|\, p)\mathrm{d}X\right) = \frac{\partial^2 p}{\partial \boldsymbol{\theta}^2}\times\frac{\partial}{\partial p}(1) = 0.$$

So the second part of the term over which we take the expectation to obtain the parameter-focused Fisher information disappears, leaving only the following:

$$\mathcal{I}(\boldsymbol{\theta}) = -\mathbb{E}^X\left[\frac{\partial^2}{\partial p^2}\ln \ell(X \,|\, p)\left(\frac{\partial p}{\partial \boldsymbol{\theta}}\right)^2\right] = -\left(\frac{\partial p}{\partial \boldsymbol{\theta}}\right)^2\mathbb{E}^X\left[\frac{\partial^2}{\partial p^2}\ln \ell(X \,|\, p)\right] = \mathcal{I}(p)\left(\frac{\partial p}{\partial \boldsymbol{\theta}}\right)^2.$$

The base level Fisher information for the Bernoulli process is

$$\mathcal{I}(\boldsymbol{\theta}) = \left(\frac{\partial p}{\partial \boldsymbol{\theta}}\right)^\dagger \frac{1}{p}\left(\frac{\partial p}{\partial \boldsymbol{\theta}}\right),$$

but with small probabilities (such as those in CKMR models) this can scale very poorly numerically. Consider $f = \sqrt{p}$, then

$$\frac{df}{d\boldsymbol{\theta}} = \frac{1}{2\sqrt{p}}\frac{dp}{d\boldsymbol{\theta}},$$

and

$$\left(\frac{df}{d\boldsymbol{\theta}}\right)^\dagger\left(\frac{df}{d\boldsymbol{\theta}}\right) = \frac{1}{4p}\left(\frac{dp}{d\boldsymbol{\theta}}\right)^\dagger\left(\frac{dp}{d\boldsymbol{\theta}}\right).$$

Given this it follows that:

$$\frac{1}{p}\left(\frac{dp}{d\boldsymbol{\theta}}\right)^\dagger\left(\frac{dp}{d\boldsymbol{\theta}}\right) = 4\left(\frac{d\sqrt{p}}{d\boldsymbol{\theta}}\right)^\dagger\left(\frac{d\sqrt{p}}{d\boldsymbol{\theta}}\right),$$

which is more numerically stable and easier to deal with.

Figures

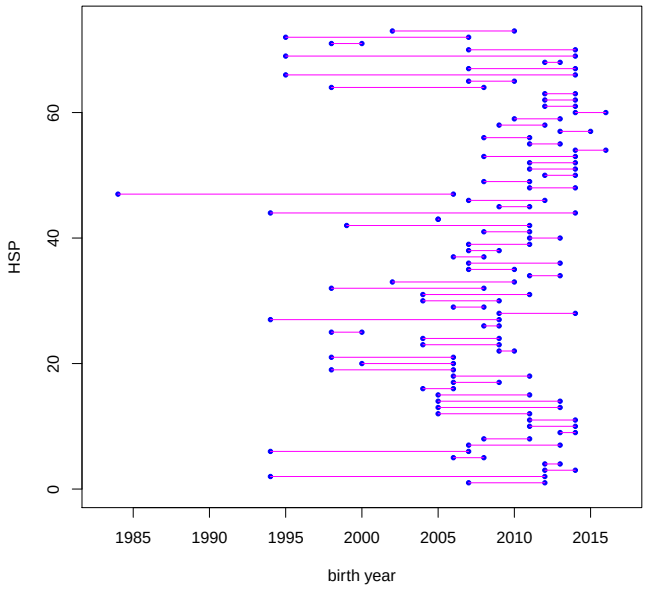

Figure 1: *Time between birth years for the 73 HSPs used in the Eastern population CKMR analyses.*

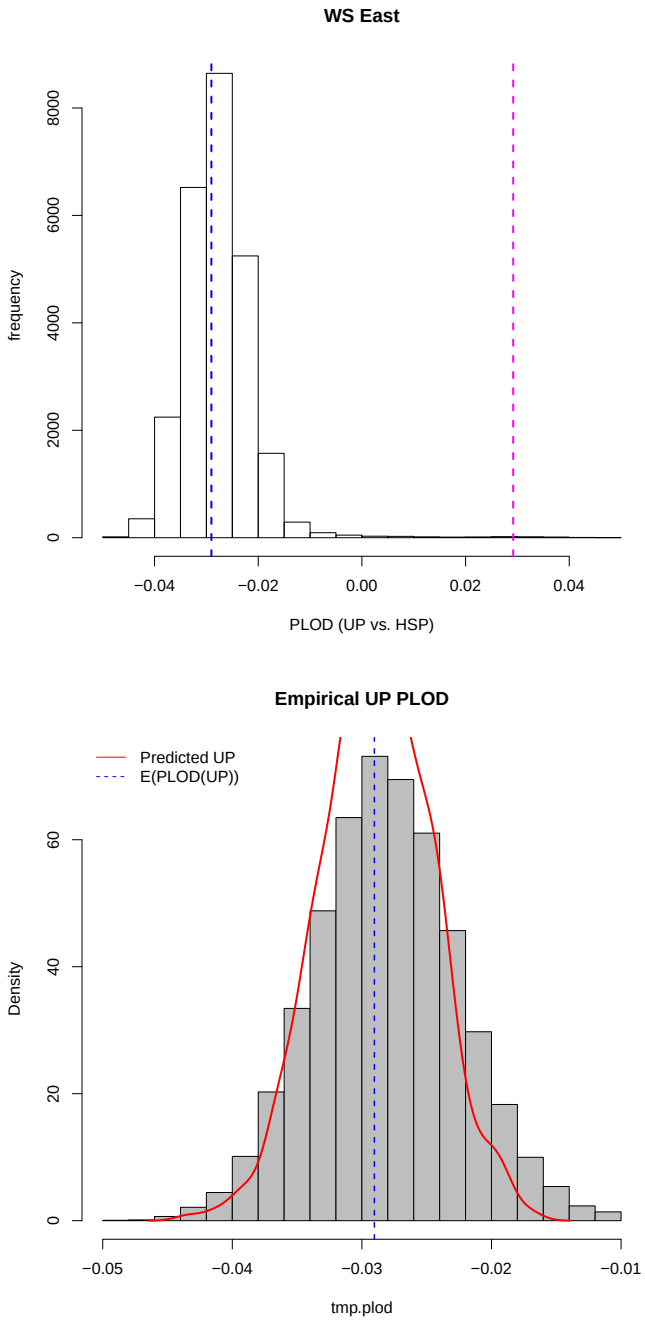

Figure 2: *PLOD histogram for UP vs. HSP comparisons (top) with UPs on the left (and expected PLOD the dotted blue line) and the HSPs (and expected PLOD in magenta) on the right. Below compares the empirical UP distribution, the predicted mean (dotted blue line), and predictive distribution in red.*

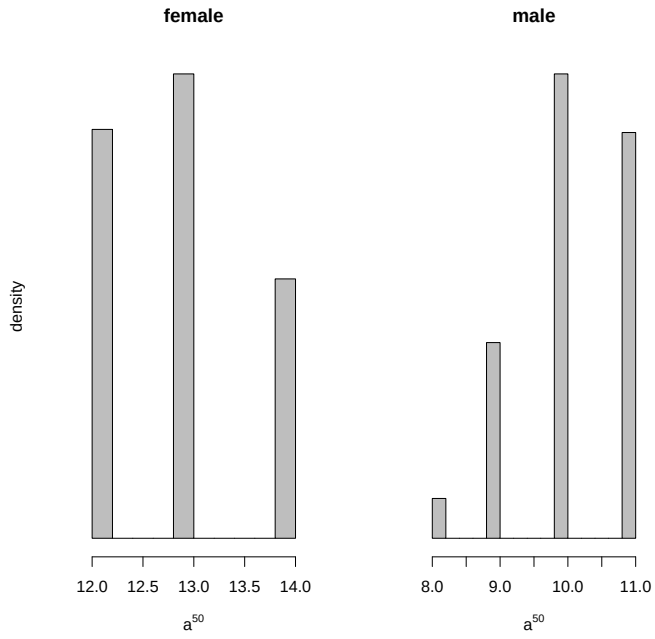

Figure 3: *Posterior summaries for female (left) and male (right) age-at-50% maturity.*

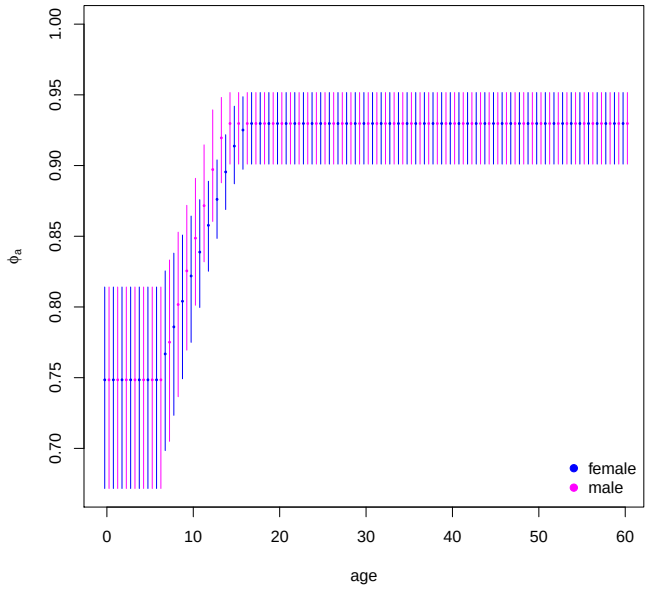

Figure 4: *Posterior summaries (median and 80%CI) for female (blue) and male (magenta) survival probability-at-age in the Eastern population.*

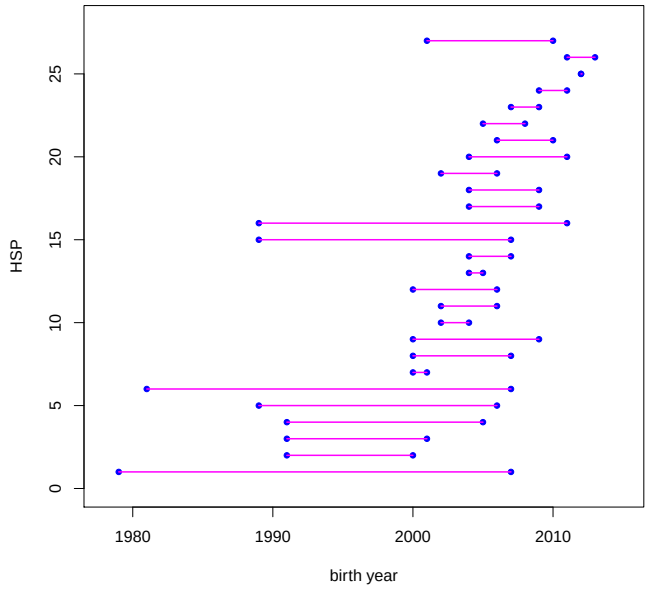

Figure 5: *Time between birth years for the 27 HSPs used in the Western population CKMR analyses.*

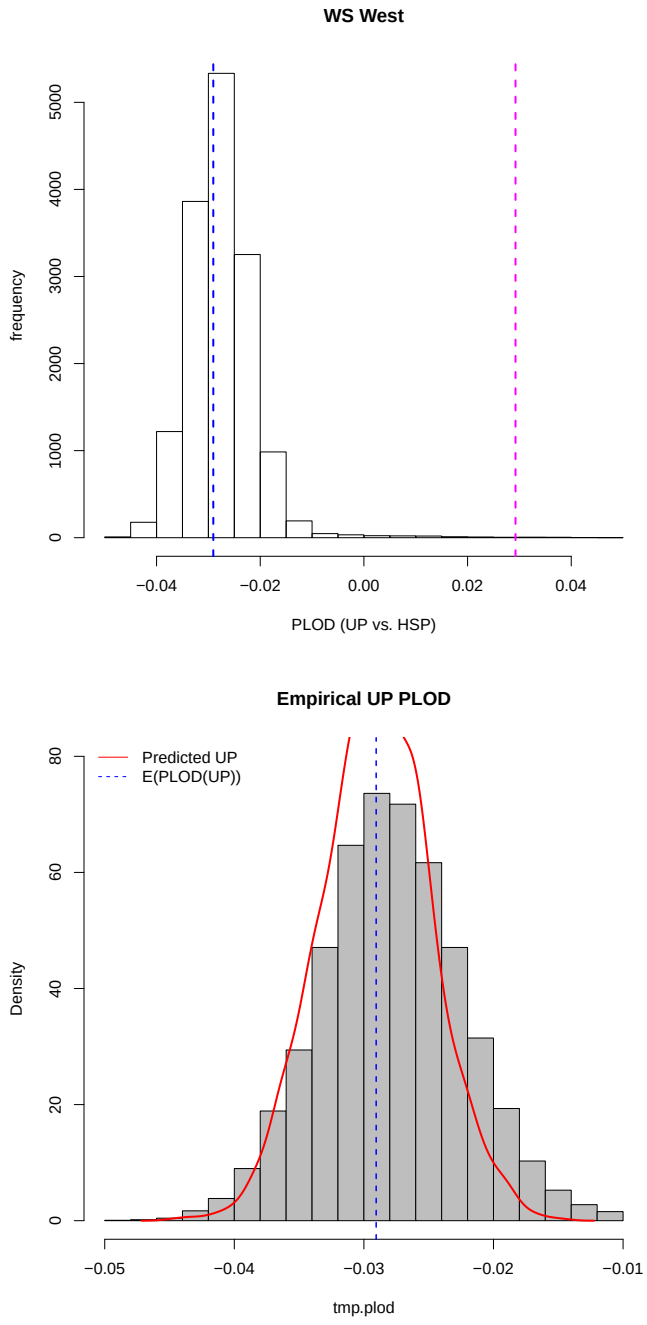

Figure 6: *PLOD histogram for UP vs. HSP comparisons (top) with UPs on the left (and expected PLOD the dotted blue line) and the HSPs (and expected PLOD in magenta) on the right. Below compares the empirical UP distribution, the predicted mean (dotted blue line), and predictive distribution in red.*

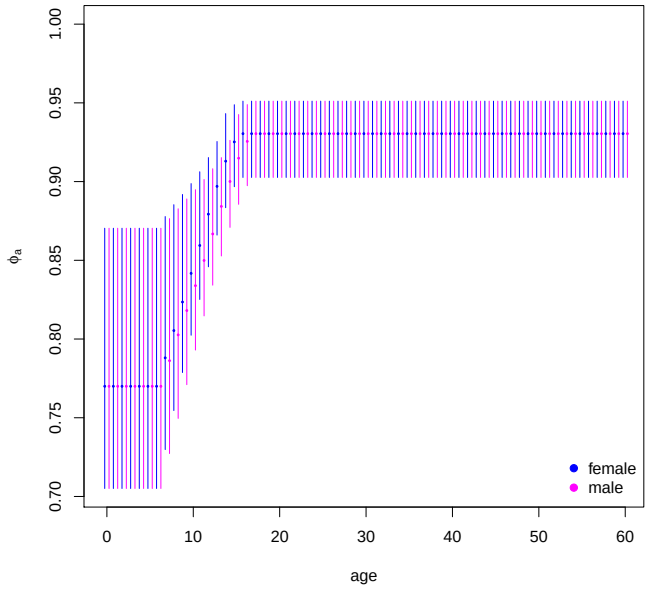

Figure 7: *Posterior summaries (median and 80%CI) for female (blue) and male (magenta) survival probability-at-age in the Western population.*

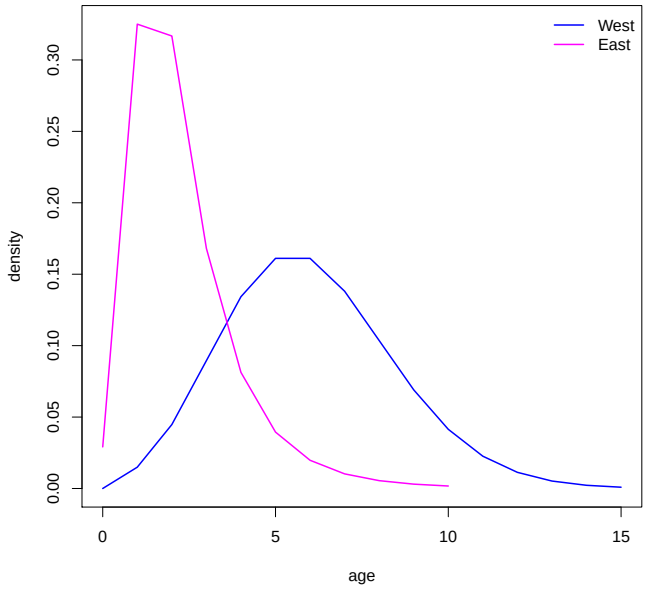

Figure 8: *Sample age distribution for the Eastern (magenta) and Western (blue) populations used in the design work.*

www.nespmarine.edu.au

Contact:
Barry Bruce / Rich Hillary
CSIRO Oceans & Atmosphere

GPO Box 1538, Hobart, Tasmania, 7001

tel | +61 3 6232 5222