Desk Review of Queensland shark stock assessment for Fisheries Queensland

Enric Cortés

Panama City, Florida, USA

February 2016

Table of Contents

1. Executive Summary 3

2. Summary of findings 6

    2.1. Data availability 6

        2.1.1. Fishery data sources 6

        2.1.2. Catches 7

        2.1.3. Species included, biology, and regional structure 9

    2.2. Demographic analysis 12

    2.3. Commercial catch rate standardization 15

    2.4. Population dynamics model 17

    2.5. Stock assessment results 22

    2.6. Management implications 24

    2.7. General conclusions and recommendations 25

3. References 28

Appendix 1 Terms of Reference 30

1. Executive Summary

An independent desk review of the document entitled "Stock assessment of whaler and hammerhead sharks (Carcharhinidae and Sphyrnidae) in Queensland" commissioned by the Queensland's Department of Agriculture and Fisheries was undertaken during January 15-February 12, 2016. The terms of reference (ToRs) for this peer review are included in Appendix 1. This review addresses the points specified in the ToRs as well as other items that are believed to be relevant, but not in the same order as in the ToRs. The review is structured by first presenting an overview/summary of each relevant item, followed by detailed findings. The items are not always presented in the same order as in the assessment to facilitate the flow of the review.

The stock assessment of whaler and hammerhead sharks represents a comprehensive examination of the status of a suite of shark species in northeastern Australia. A lot of work was put into this well written assessment, the main limitation of which were both the quantity and quality of data available. The major data limitations stemmed from the fact that the logbook data only started identifying sharks to species level in 2003 and even then identification was too unreliable to allow for species-specific assessments to be undertaken. The species composition of the catch complex had to be approximated using data from a Fishery Observer Program (FOP). Even the total catches reported in the logbook program are uncertain because they do not include discards. The assessment thus had to heavily rely on the FOP, which was conducted in 2006-2012, and, while voluntary, provided a snapshot of the gillnet fishery removals both in terms of region-specific species compositions and length compositions. One of the major limiting factors for the credibility of the assessment results was the lack of species-specific indices of relative abundance. Catch-per-unit-effort (CPUE) was computed for the shark assemblage in each of 10 subregions, but the species composition in each of these subregions was unknown (other than inferred from limited observer program data) and likely varied interannually, making it almost impossible to understand what the CPUE series are tracking in each subregion. This is due to the fact that trends for some species could be masked by opposing trends from other species in the shark assemblage with different productivity. Adding to the difficulty of interpreting the CPUE trends was the fact that no credible series of fishing effort was available.

The assessment took advantage of the relatively data-rich biological information available to develop a modified demographic method based on assumptions about natural mortality of juveniles and adults under density-dependent considerations that allowed for the reparameterization of the Beverton-Holt stock-recruit curve in terms of the recruitment compensation ratio (denoted as r_{lim}). The advantage of this metric is that it can be calculated based on life history variables alone and can then be imputed into the stock assessment model as a fixed parameter. This new method is interesting but should be further explored because the range of r_{lim} values obtained was very narrow (2.22 to 3.02), corresponding to steepness values of 0.36-0.43, considering the significantly different life histories of the thirteeen small, medium, and large whaler and hammerhead shark species groups included in the analysis. Assumptions on the ratio of the fishing mortality rate to natural mortality rate (F/M=1) as a reference point for sustainable fishing and the ratio of parental stock size with respect to virgin levels (B/B_0=0.2) at which maximum compensation from fishing would occur are shown to yield productivity estimates that are likely too optimistic based on recent findings for these quantities for shark species. Although the main conclusions from the assessment will likely not be affected, a sensitivity analysis incorporating more widely accepted assumptions about these ratios (e.g., F/M=0.5 and B/B_0=0.20-0.50) is recommended to evaluate the effect on population status metrics (e.g., $B_{current}/B_{virgin}$, MSY, F_{MSY}) obtained when fixing r_{lim} at lower values in the stock assessment.

The stock assessment model used was a length-based, sex- and age-structured model that was aggregated over 61 populations, defined as combinations of one of 10 subregions and one of 12 shark species groups. The model was modified from ADMB code used for cabezon and was fitted to catch rates from the logbooks, length compositions from the observer program, and species composition data also inferred from the observer program. The model configuration assumed virgin conditions in 1974, which seems reasonable based on the historical description of the fisheries operating in the area.The estimated parameters were virgin recruitment (R_0) and several parameters describing both logistic and dome-shaped vulnerability curves, which is fairly standard for these types of models applied to sharks. In this case, annual recruitment deviations were not estimated because the available length compositions were constant, i.e., not available on an annual basis. Maximum sustainable yield (MSY) was calculated as a separate step after model fitting using a slightly modified method that made use of r_{lim}.

Assessment results were very uncertain, showing a very large range of variation in estimated population sizes, especially at the upper end for which population sizes were deemed to be unrealistically high. This led the author to adopt a precautionary approach and focus on the results that related to the lower limit of total MSY (MSY summed over all populations) and select two deterministic scenarios: 1) a “substitute maximum likelihood point” which was on the order of 5,000 t, or about an order of magnitude smaller than the maximum likelihood estimate deemed to be unrealistic, and 2) a “minimum MSY point” which was ca. 1,300 t.

There are several aspects of the model results that are concerning. First, the model was unable to fit the abundance indices. In general terms, this was probably the result of the conflict between the catches, which showed an increasing trend from 1974 to 2003, followed by a decline from 2003 to 2013 and the CPUE series, which showed increasing (5), flat (2), or decreasing (3) tendencies from 1992/93 to 2013. The model thus interpreted that there had been very little depletion (range of $B_{current}/B_{virgin}$=0.82-0.99 and 0.44-0.96, but most values>0.75, for scenarios 1 and 2, respectively). So even in these two highly conservative scenarios, there was very little depletion from virgin levels and the model interpreted that the populations can sustain much higher removals (MSY) because of the generally increasing catch rates and reduced catches in the past decade. The problem is that the catches are likely to be underreported, but assumed to be known without error, and the CPUEs for the ten subregions do not represent the relative abundance of any species in particular, casting serious doubt on the results of the assessment.

Although results of the assessment do not show the fits to the time-aggregated length compositions, just the estimated vulnerability curves, several of the curves, especially the dome-shaped vulnerabilities, showed predicted density below the length at birth, presumably indicating that the fits to the length compositions were not satisfactory and affecting the age-based vulnerabilities used in the population dynamics model after back-transforming lengths into ages internally through the von Bertalanffy growth curve.

The assessment did not include any reference points that explicitly referred to an “overfished” or “overfishing” condition, just depletion from virgin levels (B/B_0) and F_{MSY}. However, based on MSY values for the two deterministic scenarios considered to be more credible (3,790 and 1,077 t), the current total allowable commercial catch (TACC) of 600 t and current catches of 237 t in 2013 for the Queensland east coast were thus considered to be well below reasonable values of MSY. The MSY estimate for the Gulf with the most conservative scenario (196 t) was only about 10% below the reported 2013 catch of 221 t. Given the uncertainty on the status of the 61 populations and the limitation of the data, no projections of future stock condition under alternative harvesting strategies were undertaken.

In conclusion, the assessment used the data available to the fullest extent, yet results were severely hampered by data limitations, in particular an understanding of the derived catch rates and the uncertainty about the magnitude of the catches. While a number of sensitivity analyses addressing several methodological, modeling, and other issues could be undertaken, the main data limitations that preclude the assessment to be more credible will persist until species identification improves. Future assessments of Queensland shark resources with the current data limitations will continue to yield very uncertain results and ineffective management advice and thus is not recommended. Rather than including a long list of items that may never get accomplished because of lack of funding and other issues, this review proposes two main recommendations. If the management authority is genuinely interested in assessing the status of shark resources, there has to be a serious investment in data collection. It is the opinion of this reviewer that the most effective way to collect the additional information required to conduct an improved and more credible assessment of shark resources in the area is to resurrect the Fishery Observer Program.This would provide crucial pieces of information on the annual catch of gillnet fisheries: species compositions, length compositions, status and fate of captured animals, development of an alternative relative abundance index, and collection of biological samples for life history and genetic studies. Alternatively, a fishery-independent survey of shark resources could be initiated covering the entire area and with similar objectives to those above, but this would likely require collaboration from academic partners and standardization of methodology. If distribution of field guides (for whole animals or carcasses) is thought not to be an effective strategy for species idetification by fishers or dock-side personnel, another relatively low-cost strategy could be to take fin clips from specimens captured in the gillnet fishery to genetically identify them and then compare the species ID to that reported in the logbooks. This could be a way to verify the inaccuracy of the species composition in the logbook data, attempt to reconstruct the species composition back in time, and adjust newly acquired logbook data.

2. Summary of findings

2.1. Data availability

2.1.1. Fishery data sources

<u>Overview</u>

There were three data sources initially available from: 1) the commercial logbook system, 2) the Fishery Oberver Program (FOP), and 3) the Shark Control Program (SCP). Data from commercial logbooks (1988-2013) only identified sharks to species level starting in 2003 and even then were deemed too unreliable to establish species compositions due to the inaccuracy of species identification by fishers and thus were not used in the assessment for that purpose. However, they were used to quantify total catches and to calculate catch rates for aggregated shark species by subregion. Data from the FOP (2006-2012), while voluntary, provided a more reliable snapshot of species composition of the catches as well as species-specific length frequencies and information on the fate of sharks caught. Use of data from the SCP was explored, but abandoned because of numerous changes in the gear used and concerns with species identification.

<u>Detailed findings</u>

I find that the *non-use of logbook data to establish species compositions (ToR1c)* is justified because in general sharks are notoriously difficult to identify and multiple reasons that could have affected the accurate reporting to species by fishers. As recognized in the assessment, some of the concerns affecting species identification in the logbooks could also have affected the total catch reported and the calculated catch rates. Some of the main reasons include changes in target species, which could differentially affect the species of sharks caught as well as the non-reporting of discarded sharks and varying incentives to accurately report catches. It is interesting to note that the beginning of species identification in the logbooks in 2003 coincides with the onset of a decline in total reported catches and suggests that catches may have been underreported from 2003 onwards. In general, total catches from the logbook system are also subject to error and may be biased low due mostly to the non-reporting of discarded sharks (see Catches section 2.1.2 below).

The *use of data from the FOP for species identification, sizes of sharks, and discard rates (ToR1b*) is a consequence of the unavailability of other reliable sources and thus almost inevitable. While the dataset is not perfect because the program was voluntary, only spanned a period of 6-7 years, included a small proportion of sharks caught in other gears, and there were also some concerns about the realiability of species identification even by trained observers, it provided essential information to attempt to conduct this assessment.

The *non-use of data from the SCP* is also justified because of the multiple changes in gear configuration, problems with species identification, and other factors. It is unfortunate because the program started in 1962, twelve years before the assumed beginning of commercial shark harvest in the region, and could have provided a glimpse into the status of virgin, or close to virgin, populations. However, if the initial 15-year period of depletion reported in Appendix 1 of the assessment is real and not due to some artifact such as the occurrence of pre-existing gear on neighboring beaches as reported in the assessment, this would indicate that exploitation started well before 1974 as was assumed in the assessment.

2.1.2. Catches

Overview

Only commercial gillnet catches were included in the final harvest estimates used for the stock assessment because sharks are primarily caught in this gear in Queensland. Catches in line and trawl fisheries were considered to be low and not included. Recreational hook and line catches were also not included because most sharks are believed to be released alive and most of the sharks released alive are expected to survive. Historical commercial fisheries in Northern Australia include a Soviet trawl and gillnet fishery that appears to have operated very little in Queensland and for which there are largely no records and a Taiwanese fish-trawl and gillnet fishery which operated from 1974 to 1990. The Soviet fishery thus was not included in the stock assessment. Catches of less than 1,000 t annually, smaller than inferred from published reports, were included for the Taiwanese fishery for 1974-1979. For the Australian fishery, a linear increase in catches was assumed from 0 t in 1973 to the 1988 value from the logbooks.

Detailed findings

Catches in the line fishery on the east coast are believed to be of reef-dwelling sharks mostly based on the FOP and thus were not included in the total removals since the reef sharks were not even on the initial list of species to be assessed. Since the magnitude of the line fishery is believed to be small compared to the gillnet fishery, the decision seems reasonable. Similarly, catches in the prawn and trawl fisheries were not quantified, but were believed to be small because of the introduction of Turtle Excluder Devices (TEDs). The assessment does not mention in what year TEDs were introduced and so it is possible that large catches of sharks occurred prior to their introduction. The assessment also mentions a relatively large proportion of sharks released alive (23%), which is surprising if compared for example to bottom-trawl fisheries in the USA where most shark species die, although this could be due to different handling and sorting procedures of the catch. For comparison, the magnitude of estimated prawn trawl catches that die can potentially be very large and in fact dominate total catches for some small coastal shark species in the USA with similar biology to those that occur in Queensland waters (NMFS 2013). Estimation of bycatch in these fisheries is notoriously problematic due to low observer coverage, but its magnitude can be very large as noted. I assume that estimation of removals in these fisheries may not be possible in this case due to the very limited observer coverage (in time and as a proportion of total effort observed), but it should be mentioned here in case it were feasible to estimate total bycatch with catch rates from the FOP and effort data from the trawl fisheries if only to provide a rough estimate of the magnitude of the potential removals in years when TEDs were not in use.

Recreational surveys surprisingly only seem to be available for sporadic years and thus it is not possible to provide estimates of shark catches in the recreational hook and line fishery in the region. Recreational shark catches, however, can potentially be very large (as is the case in the USA). In the case of Queensland, catches are relatively small and the proportion of sharks retained in 2000-01 was only ca. 11% (or 37,000 sharks). Since sharks caught tend to be small, an assumed mean weight of 2 kg/shark would result in a catch of about 75 t. It was assumed that the majority of the sharks relased alive (ca. 89%) would survive. This is a reasonable assumption as post-release mortality in hook and line fisheries has been estimated at 10% for Atlantic sharpnose and blacktip sharks and even lower rates for other species (Courtney 2013). Nevertheless, in the example above, assuming 251,000 sharks (288,000-37,000) were released alive and 10% died would add another 50 t (25,100x2 kg) to the estimated total. This is still a small magnitude when compared to the commercial catches, but I mention it because recreational catches can potentially be larger than reported.

The next paragraphs refer to the *estimated historical harvest levels going back to the mid-1970s (ToR1d)*. The non inclusion of catches from the Soviet trawl and gillnet fishery that operated in Northern Australia and only partially in Queensland is justified based on the lack of records and the little overlap with Queensland. The Taiwanese gillnet and fish-trawl fishery operated off Northern Australia from 1974 to 1990, but effort after the declaration of the Australian Fishing Zone in 1979 was believed to be negligible for that area. By comparing official and unofficial records of catches of sharks and fish combined, using a proportion of 78% of sharks derived from Walter (1981), and a 10:7 factor to convert processed weight to whole weight, catches of sharks were estimated for 1974-1982 and the higher unofficial catches deemed more reliable. Catches in 1974-1979 from records in a logbook database maintained by ABARES were then presented to show that sharks as a group only made up ca. 20% of the total reported catch and that the estimate of 78% reported by Walter (1981) was an overestimate. Catches in the Queensland Gulf of Carpentaria from the above database were then tabulated and scaled up to match the unoffocial catches for all shark and fish species mentioned above for years in which the total logbook harvest was less than the unofficial harvest. The resulting final catches from the Taiwanese fishery for the Gulf ranged from 0.2 to 812 t from 1974 to 1979. These estimates seem more justifiable than those that would be obtained by using the anecdotal proportion of 78% obtained from the fishery, which resulted in estimates of much higher magnitude. I note that in Figure 14, the catches imputed into the stock assesment for the Gulf in 1977 for example seem to be ca. 550 t whereas those listed in Table 18 for the same year are 812 t.

For the Australian fishery, the assessment assumed a linear increase in catches from 0 t in 1973 to the 1988 value reported in the logbooks. The justification for this assumption was that the logbook catches did not trend sharply upwards beginning in 1988, thus suggesting that they had been sustained for a number of years, that the 1980-1981 logbook catches may have been underreported, but that changes in net technology that would have allowed increased capture of sharks did not take place until the 1970s, suggesting that shark catches in gillnets prior to the mid-1970s were probably much lower than in the late 1980s. In the absence of more detailed information, these assumptions seem reasonable, although they do not rule out the possibility that there still may have been shark catches in gillnets prior to that period. This could affect the results of the stock assessment since the population is assumed to be at equilibrium (unexploited) at the beginning of the model year, but some depletion from virgin levels may already have occured. In conclusion, catches for 1974-1987 were estimated from several sources, and catches for 1988-2013 came from the logbooks, in all cases for all sharks combined.

2.1.3. Species included, biology, and regional structure

<u>Overview</u>

The original target species for the assessment included five carcharhinid sharks and one sphyrnid shark, which were classified into four categories on the basis of their length (small whaler, medium whaler, large whaler, and hammerhead) for convenience. Six additional species groups, consisting of 15 additional species, commonly caught in gillnet fisheries had to be added to the list because of concerns with incorrect species identification in the fishery data. The resulting 12 species groups comprised a total of 21 species. Additionally, less common species were also allocated to these 12 groups based on similarity in biology. In the end, more than 40 species were included in the 12 species groups that were used in the stock assessment.

The re-parameterized von Bertalanffy growth function expressed with length at birth (L_0) rather than t_0 was used to describe length at age. Although growth dynamics of males and females typically differ in sharks, the assessment did not model male and female growth separately for reasons of complexity and due to lack of quantitative data for some species. The assessment also assumed that catches are composed of males and females in equal proportions.

The demographic model (see Demographic analysis section 2.2) assumed females had a constant litter size because of a lack of quantitative data, whereas the stock assessment population model assumed fecundity was proportional to weight. The assessment model also assumed a 1:1 sex ratio at birth. Life history inputs summarized for the 12 species groups considered (13 species) were the coefficient and slope of the weight-length relationship, mean length at birth (L_0), k and L_∞ from the von Bertalanffy growth curve for males and females separately, lifespan (maximum observed age), the ages at 50% and 95% maturity of females, pupping interval, and litter size.

The stock assessment model was structured regionally, based on sampling regions used by the Queensland's Long Term Monitoring Program (LMTP). Although 22 regions are identified in the LMTP, only 19 of them had catches big enough to define meaningful catch rate time series and species composition from the FOP. These regions were further merged into a total of 10 subregions to reduce statistical error in the standardized catch rates. The final number of subregions used in the stock assessment model was thus 10.

Detailed findings

Addition of the six species groups to the original six target species is justified because of the uncertainty in species identification. The further allocation of less common species to these 12 species groups based on biological similarity likely had very little impact since the 12 species groups account for the vast majority of catches according to the FOP. However, it must be noted that the species included in these categories, which range from small to medium to large whalers and hammerheads, have substantially different life history strategies, which will affect their productivity.

The assessment used the von Bertalanffy growth function with fixed L_0, rather than the formulation that uses t_0, which means that only two parameters (k and L_∞) were estimated. This is reasonable and more biologically realistic in my opinion, but some have recently argued that the L_0 parameter should be estimated on the grounds that it can result in biased growth estimates (see Pardo et al. 2013). Although sex-specific growth function parameters were available for the 13 main species (Table 6 of assessment), it appears that only values for females were used in the assessment model. Given the poor quality of data available for the assessment and the large number of assumptions that had to be made regarding species compositions, using a true sex-specific model may have resulted in a false sense of precision and knowledge and thus I do not consider it a flaw.

In terms of reproduction, while there is a trend of increasing litter size with increasing maternal size in many shark populations, this is not a universal trait in sharks (see e.g., Cortés 2000). Although it is mentioned that the demographic model used a constant litter size because of the lack of quantitative data, I wonder if no litter size vs. maternal size relationship was available for any of the 13 major species (it may be the case for Australian populations). The assumption of a 1:1 sex ratio at birth is justified as this is generally what is reported in most published studies. Of the other life history inputs listed in Tables 6 and 7, I wonder what the advantage was in using this specific formulation of the maturity ogive (with a_{50} and a_{95}). Since age at maturity was derived from length at maturity, a_{50} was obtained from L_{50} by back-transforming through the von Bertalanffy growth function and a_{95} was arbitrarily set to a value $\geq a_{50}$ depending on the value of a_{50} (Table 7). This is probably because the cabezon code only allows for a functional form of proportion maturity at age to be specified and not a parameter vector of maturity at age to be imputed. Otherwise, it would have been easier to use a length-based maturity ogive (if available) and back-transform to each age class.

The general idea of structuring the assessment into subregions was to take advantage of the information available in the LMTP and FOP and obtain a more detailed depiction of shark relative abundance in Queensland given the dramatic changes in relative abundance that can occur over a small spatial scale. This aggregation allowed the calculation of 10 indices of abundance, obviating the need to include “region” as a factor in the catch rate standardization (see Commercial catch rate standardization section 2.3). These 10 subregions were later used in combination with the 12 species groups identified earlier to define 61 populations (subregion-species group pairs).

2.2. Demographic analysis

Overview

A *new demographic analysis (ToR1a)* that builds on methodology by Smith et al. (1998) was introduced. The method has the advantage of expressing results in terms of stock-recruitment parameters that are more amenable for use in stock assessments as well as providing estimates of natural mortality that can also be directly imputed into the stock assessment as fixed parameters. Briefly, the method assumes 1) that the population is in virgin conditions, with no fishing, and estimates a value of M (natural mortality rate); 2) a fishing mortality of the same magnitude as M is then imposed on adults only and it is assumed that the steady state condition is maintained by a decrease in M of juveniles; and 3) demographic parameters are converted to recruitment compensation ratios (Goodyear 1977). The method also assumes that F/M=1 is a sustainable reference point (Gulland 1970). Maturity in the method is modelled with the same logistic functional form as described before and the total mortality rate ($Z=M+F$) takes one value for ages <5% maturity (Z_{juv}) and another value for ages ≥5% maturity (Z_{adult}). The method also assumes that in the first step (virgin conditions) the ratio M_{juv}/M_{adult} is equal to a fixed value, which is taken as the minimum value that satisfies two conditions: 1) that M_{prod}/M_{adult} ≥1.5 and 2) that the proportion of sharks above the maximum observed age (a_{max}) must be at least 0.002. M_{adult} is then solved for from their equation (2.1) after fixing M_{juv}/M_{adult}. In the second step (exploited conditions), it is assumed that Z_{adult}=2 M_{adult} and Z_{juv} is solved for from equation (2.1); the productivity-adjusted value is denoted by M_{prod}, which is always larger than M_{adult} and represents the recruitment compensation that would occur after removing fishing but retaining the lower juvenile natural mortality rate. The Beverton-Holt stock-recruit relationship is then expressed in terms of this recruitment compensation ratio, which is defined as the "average number of animals produced by each animal over its lifetime at extremely low population sizes". A final assumption is that maximal recruitment compensation takes effect at some parental population size B_{prod} such that B_{prod}/B_0=0.2 (maximum recruitment compensation occurs at a stock size of 20% of virgin level).

Results of the demographic analysis showed that the recruitment compensation ratio (r_{lim}) ranged from 2.22 to 3.02, or expressed as steepness, from 0.36 to 0.43.

Detailed findings

I found the demographic method presented was an interesting modification of the density-dependent method first introduced by Smith et al. (1998), with the advantage that it estimates natural mortality internally and that both these natural mortality rates and the recruitment compensation ratios can be imputed as fixed parameters in stock assessments (or they could be given a prior too in a Bayesian estimation context). I have the following observations about the method:

1) A similar metod was developed by Brooks et al. (2010 and related papers) based also on Goodyear's (1977) compensation ratio and Maximum Excess Recruitment (MER; Goodyear 1980), as well as Myers et al.'s (1997) maximum lifetime reproductive rate at low population density ($\hat{\alpha}$), which is essentially the same as the recruitment compensation ratio defined in the assessment. This method in fact analytically derives species-specific values for $\hat{\alpha}$, the associated steepness, and the depletion of spawners and recruits corresponding to MER in the Beverton-Holt model, which is equivalent to the B_{prod}/B_0 ratio reported in the assessment.
2) The rule of thumb that F_{MSY}/M=1 from Gulland (1970) has recently been shown not to be supported by empirical data. Au et al. (2008), who are the same authors as in Smith et al. (1998), concluded that F_{MSY}/M =0.5 and Zhou et al. (2012) found that F_{MSY}/M=0.41 for chondrichthyans. These new rules of thumb do not take account of selectivity and so these ratios may be even lower for some species of sharks.
3) Equation (2.1) in the assessment defines S_a as the probability that an animal survives to the end of year *a*, but it should be the cumulative probability.
4) Gestation period should be incorporated as was done by the author. It should be the age at first breeding that is considered, not the age at maturity.
5) Total mortality, Z_i: what is the rationale for Z_{juv} being defined for ages less than the age at 5% maturity? Why this very low number in particular?
6) M_{juv}/M_{adult} ratio: why does it have to satisfy the rather arbitrary condition of M_{prod}/M_{adult}>=1.5?
7) For the fished population, Z_{adult}=2*Madult*. As noted above, should be Z=1.5M.
8) The statement that Smith et al. (1998) use S_{a50} as the juvenile mortality parameter but that they allow M_{juv}<M_{adult}: Smith et al. (1998) defined cumulative survival to age at maturity ($l_{\alpha,2M}$) or $l_{\alpha,1.5M}$ (Au et al. 2008), not M_{juv} for any given juvenile year. If a constant M is used for juvenile ages, l_α always produces M_{juv}>M_{adult}. This can easily be checked by getting M_{juv} from –ln(l_α)/a_{50} and comparing it to the M values for adults listed in Table 3 of Smith et al. (1998).
9) B_{prod}/B_0=0.2: as mentioned above, the method by Brooks et al. (2010) produces species-specific values with an upper limit of 0.5, with values for less productive species probably approaching the upper limit. Sainsbury (2008) also advocated a limit reference point for sharks of 0.3B_0.

I found that the results of the demographic analysis showed a very narrow range of recruitment compensation ratios (or steepness) considering the substantially different life histories exhibited by the 12 species groups included in the analysis. This lack of contrast is concerning. Since the demographic analysis constrained M_{prod}/M_{adult} or fixed several important quantities (*F/M*, M_{juv}/M_{adult}, B/B_0) at values that may not be the most appropriate for sharks, I recommend conducting an *extensive sensitivity analysis (ToR3)* to assess the influence of these assumptions on obtained quantities. I conducted a brief exploratory analysis with three of the species groups (Australian sharpnose, Australian blacktip, and great hammerhead) using the R code provided by the author and found that using a value of *F/M*=0.5 and a value of B/B_0 other than 0.2 yielded lower estimates of r_{lim} and steepness, as expected (see Table 1 below). Lower values of productivity will affect the results of the stock assessment and likely show more depletion than was obtained (see Stock assessment results section 2.5).

Table 1. Comparison of results from the demographic analysis with those obtained by the reviewer for three species groups using the R code provided by the author and showing the effect of different assumptions on the ratios of *F/M* and B/B_0 on calculated quantities. The results for *R. taylori* could not be exactly replicated using the parameter values listed in the assessment document (yellow highlights). Cells highlighted in gray denote changes introduced to the *F/M* scalar and the B/B_0 ratio.

Inputs										Reviewer results							Assessment results				
Species	a50	a95	maxage	PupFreq	LS	Mjuv/Mad	F/M	B/B0	I-PLUS	M1	M0	r_lim	r	surv	Mprod/Madult	steepness	M0	r_lim	r	Mprod/Madult	steepness
R. taylori	1	1	6	1	4.5	3.78	1	0.2	max age	0.649	0.386	2.249	3.271	0.005	1.680	0.36					
R. taylori	1	1	6	1	4.5	3.57	1	0.2	max age	0.601	0.400	2.284	3.363	0.005	1.503	0.36	0.401	2.494	3.980	1.503	0.38
R. taylori	1	1	6	1	4.5	3.78	*0.5*	*0.25*	max age	1.022	0.386	1.549	1.895	0.012	2.647	0.28					
R. taylori	1	1	6	1	4.5	3.57	*0.5*	*0.25*	plus group	0.961	0.414	1.678	2.169	0.022	2.320	0.30					
C. tilstoni	5.65	6.7	25	1	3	2.74	1	0.2	plus group	0.226	0.150	2.529	4.095	0.010	1.504	0.39	0.150	2.529	4.095	1.504	0.39
C. tilstoni	5.65	6.7	25	1	3	2.74	*0.5*	*0.3*	max age	0.306	0.139	1.456	1.809	0.006	2.200	0.27					
C. tilstoni	5.65	6.7	25	1	3	2.74	*0.5*	*0.3*	plus group	0.307	0.150	1.689	2.395	0.023	2.042	0.30					
S. mokarran	7	8	39	2	15.4	17.92	1	0.2	plus group	0.799	0.052	2.216	3.183	0.002	15.365	0.36	0.052	2.216	3.183	15.365	0.36
S. mokarran	7	8	39	2	15.4	17.92	*0.5*	*0.4*	max age	0.847	0.050	1.389	1.874	0.0004	16.834	0.26					
S. mokarran	7	8	39	2	15.4	17.92	*0.5*	*0.4*	plus group	0.855	0.052	1.579	2.572	0.0048	16.451	0.28					

2.3 Commercial catch rate standardization (ToR1e)

Overview

Catch rates were calculated using data from the commercial logbook system, which spanned 1988-2013. Since species identification was very unreliable, catch rates of sharks in adjacent regions displayed very different trends, and data from the FOP also showed that species composition varied greatly over small spatial scales, catch rates were computed for the 10 subregions identified earlier for all species combined. Other important caveats were that: 1) only sharks retained were reported, not those discarded, 2) it was initially assumed that the major alternative target species was grey mackerel and thus only records containing sharks or grey mackerel were extracted, 3) a management change was introduced in 2009 whereby fishers with an “S” license symbol were allowed to retain 10 or more sharks per set, 4) the analysis was performed on positive catches only, i.e., fishing days that reported at least one shark being caught, the reason being that when zero catches of sharks were reported (but grey mackerel was reported) this likely indicated that sharks had been discarded, 5) some catch rates for specific years and regions were excluded after the standardization was performed because of suspected high discard rates in those years, 6) the nominal catch rate was weight of sharks retained per fishing day (if multiple sets in different locations took place in one day, the set with the maximum catch weight was used), 7) catch rates for 1988-1990 were not used because of inconsistent reporting of catches in those years, 8) net depth was not recorded in the logbooks, and 9) it is unclear if net length represents that actually used or that allowed.

The Generalized Linear Model (GLM) used for catch rate standardization used a quasi-Poisson error distribution and the explanatory variables included were year, month, boat (a fisher identifier), and net length. The “boat” variable was found to have a disproportionately high influence on results in all cases. The assessment concluded that, geographically clockwise, CPUE showed no trend in the Gulf, Far North, and Lucinda subregions, declines in Whitsunday and Stanage, but increasing tendencies in the two neighboring Rockhampton subregions, no trend in the Fraser inshore subregion, and an upward trend in the Sunshine Coast Offshore and Moreton subregions.

Detailed findings

Computation of catch rates for 10 subregions is justified in order to take advantage of the information from the logbooks. This also obviated the need to include “area” as one of the explanatory variables, which would undoubtedly have explained a large proportion of the variance in that case. Addressing the caveats in the same order in which they were presented in the overview, I conclude that: 1) although the non inclusion of discarded sharks in the logbooks is obviously problematic, discard rates, as inferred by the proportion of nonzero shark catches, appear to have been pretty consistent in most regions and thus not overly affected results; 2) it is unclear what the implications of assuming that the alternative target species was only grey mackerel were: inclusion of additional records with other target species may affect the derived catch rates; 3) the introduction of the “S” symbol does not seem to have biased discard rates, as concluded in the assessment; 4) limiting the analysis to the positive catch rates is a reasonable approach in this case; 5) eliminating values for specific years in each subregion due to concerns with high bycatch rates is also reasonable, but I wonder why this was done after the standardization and not before; 6) I also wonder why the nominal catch rate used was fishing day rather than set. Also, effort, expressed as the number of hours fished, could have been used as an offset in the Poisson regression, but as the author indicates it was not used because of the unreliability of that measure; 7) exclusion of 1988-1990 in all cases is also justified due to their unreliability, but as before why do it after the standardization?; 8) and 9), it is unfortunate that net depth is hardly ever recorded and that net length may not always be the one used as these two measures could have allowed definition of another nominal CPUE, for example, sharks caught per area of net fished.

Although there are a number of error distributions that can be used to standardize catch rates (Maunder and Punt 2004), several of them treat the proportion of zero catches and the proportion of positive catches separately (e.g., the delta-lognormal approach; Lo et al. 1992). As the standardization used the positive catches only, the Poisson regression used is adequate in my opinion and the diagnostics for the model fits appear adequate. The assessment does not explicitly mention why mesh size was not used as a factor in the standardization, but it is later mentioned (p. 71) that 71% of catches in the logbook database come from the 15-16.5 mm mesh size.

In my own interpretation of Figure 17 in the assessment, I conclude that there are five subregions that show increasing trends (Gulf, Far North, RockEst, SunshineOff, and Moreton), two subregions with flat tendencies (Lucinda and RocksOff), and three subregions with decreasing trends (Whitsunday, Stanage, and FraserIn). That the Whitsunday and Stanage subregions show decreasing trends, whereas the adjacent Rockhampton subregions show increasing or stationary trends could in theory be explained by the different species compostion in these subregions, for example with the two first subregions being dominated by less productive species that are declining in abundance and the last two subregions consisting of more productive species displaying more increasing trends. However, data from the FOP in Figure 7 reveal that the species composition for the two contiguous subregions of Stanage and RockEst are similar and thus the opposing trends in CPUE are hard to reconcile unless the relative species composition has changed with time. In contrast, the increasing trends in the SunshineOff and Moreton subregions could be explained because these two areas are composed mostly of spinners (SunshineOff) and sharpnose (Moreton) sharks and thus it is possible that the abundance of these smaller and medium sharks, especially in Moreton, is increasing.

The issues discussed above are only second-order concerns. My main concern is that it is unclear what the shark abundance in each region really represents. This is because each subregion is composed of a certain mix of species, which is not captured sufficiently by the species composition from the FOP, which is only a snapshot for 2006-2012. In reality, the species composition will have varied with time and the CPUE for each subregion encompasses a variety of trends for different species, some of which may be increasing, others stationary, and others decreasing. The CPUEs for each subregion are averaging these trends but they do not reflect the trend for any one species or even as a group because more productive species with an increasing trend in abundance may be masking the decline of less productive species. This was the situation in the 2002 assessment of large coastal sharks in the USA (NMFS 2002), which eventually led to discontinuing a stock assessment for that complex in favor of species-specific assessments.

2.4 Population dynamics model (ToR1f)

<u>Overview</u>

The population dynamics model used was a length-based, sex- and age-structured, forward-projecting model that used modified ADMB code applied to a stock assessment of cabezon to accommodate the data structure of the current assessment (multiple regions, species groups, species compositions, separate juvenile and adult natural mortality). The model was disaggregated into 10 subregions, each of which had its own catch and CPUE time series based on logbook data as well as species composition obtained from the FOP. It considered 61 populations,which represent combinations of one of the 10 subregions and one of 12 shark species groups. The model was fitted to catch rates from the logbooks, length compositions (5-cm intervals) from the FOP, and species composition data also inferred from the FOP. The model internally converts length-based vulnerabilities to age-based ones and assumes that fishing occurs as a short pulse in the middle of the year. The model assumed virgin conditions in 1974, catches covered the period 1974-2013, catch rates spanned 1992/93-2013, and species compositions and length-frequencies corresponded to those from the FOP for the combined period 2006-2012. Other assumptions of the model were that age ranged from 0 to 30 years for all species, with a “plus group” to accommodate those animals that were still alive past age 30 and that recruitment occurs at age 0. Because the catch and catch rate data were not species-specific, the model had to further assume that every species had 100% vulnerability at some length and that the same harvest rate applied to all species within a particular subregion and year.

Vulnerability—Vulnerability estimated in the model includes the selectivity of the fishing gear and the availability of the animal to the gear. Two vulnerability functions were used: a logistic function (expressed in terms of L_{50} and L_{95}), which was used for most smaller sharks; and a three-parameter dome-shaped function for larger sharks. The length-dependent vulnerability estimated by the model was then converted to sex- and age-dependent vulnerability inside the model using the distribution of length at age. Length at age was assumed to be normally distributed, with the mean obtained from the von Bertalanffy growth curve and the SD from a CV, which in theory can be estimated, but which was fixed at 0.07 for age-0 sharks and 0.05 for sharks ages 30+, and was a linear function in between.

Recruitment—Recruitment was modelled as a deterministic process. Annual recruitment deviations were not estimated because it would imply having to estimate a very large number of parameters. The stock-recruit relationship was expressed with the modified Beverton-Holt curve presented in the Demographic analysis section that makes use of the recruitment compensation ratio and the density-dependent juvenile mortality (M_{prod}). The assessment used parental stock biomass in the stock-recruit relationship and assumed no influx of recruits from neighboring areas. Mating and recruitment were assumed to occur simultaneously while in reality there is a lag between mating and recruitment. Finally, virgin recruitment (R_0) was expressed as density/unit of habitat, rather than numbers, for convenience and to reflect the different sizes of the 10 subregions considered.

Parameter estimation—Most parameters in the model were fixed. The same biological parameters used or estimated in other sections, and juvenile and adult natural mortality and the recruitment compensation ratio obtained in the demographic analysis, were used as fixed parameters in the model. The model estimated virgin recruitment ($\ln R_0$) and length-based vulnerability parameters of logistic and dome-shaped curves. Parameter estimates were obtained as usual from ADMB by inverting the Hessian matrix and uncertainty in parameter estimation was investigated with the MCMC algorithm.

Data inputs—The assessment mentions that the CVs of the standardized catch rates, expressed as the SE of the log-transformed estimate, were used in the model as lower bounds for the actual CVs (SE/mean) to account for process error. It also mentions that the observer data were considered to be collected for the year in which the most sharks were observed in each subregion.

Likelihoods—Four likelihood contributions to the objective function were specified: for relative abundance, length frequencies, species frequencies, and recruitment.

Maximum Sustainable Yield— MSY was calculated in a separate step after model fitting using a slightly modified method that made use of the recruitment compensation ratio obtained in the Demographic analysis section. Briefly, it consisted of four steps: 1) find the Yield Per Recruit (YPR) as a function of F, 2) find R/R_0 as a result of exploitation and the depleted parental stock and then multiply this ratio by YPR to obtain a Yield Per Virgin Recruit (YPVR), 3) optimizeYPVR over F, and 4) multiply the Y/R_0 ratio from the previous step by R_0 (virgin recruits) to obtain MSY.

<u>Detailed findings</u>

The population dynamics of the model are fairly standard, with the exception of the complicating factor of having to consider multiple subregions, species groups, and species compositions. Although the model was sex- and age-structured, it was only length-based since only limited length compositions were available. I believe that defining the numerous populations (61) used in the model responded to a desire to fully utilize all the available biological information and the species-specific length and species composition by subregion from the FOP. The structural assumptions of the model seem reasonable. One potential *sensitivity scenario that could be investigated (ToR3)* is the assumption that the population was in a steady state at the beginning of the model in 1974, since there may already have been some depletion from virgin levels (see Stock assessment results section 2.5). It is also unclear if *assuming the same maximum age of 30 years for all 12 species groups considered* may have affected results.

Vulnerability (ToR1g)—Considering the two functional forms of vulnerability curves, logistic for smaller species and dome-shaped for larger species, is reasonable because smaller shark species are fully vulnerable at their maximum lengths whereas larger shark species become less vulnerable after a certain peak length. I have two observations regarding the distribution of length at age. First, how were the animals in each length class apportioned into males and females since the vulnerability was sex-specific (assuming a 50/50 split)? Second, the CVs of the length-at-age distributions could not be estimated likely because the length-frequency distributions were constant, not for each year considered in the assessment. This is reasonable, but why were the values for age-0 and age-30+ sharks fixed at the particular values of 0.07 and 0.05? Furthermore, it seems counterintuitive that the CV for age-0 sharks would be larger than that for much older sharks since one would expect much less variability in length at age 0 than in length at age 30+. I suspect that fixing the CVs at these particular values may be related to the model being able to converge and thus I would recommend some *sensitivity runs to investigate the effect of the CV values(ToR3)*. At a minimum, some justification for the chosen CV values should be provided.

Recruitment—I agree with the decision of not estimating recruitment deviations because the available length compositions were constant, i.e., not available on an annual basis, and there would not have been sufficient information in the data to estimate these deviations. Additionally, recruitment deviations are expected to be low in sharks because of their more direct stock-recruit relationship. The assessment used parental stock biomass and notes that while female egg production is used for teleosts assessments, this concept does not apply to sharks. But in fact spawning stock fecundity (sum of number at age times pup production at age) is routinely used in shark stock assessments in the USA as a measure of parental stock size (see e.g., NMFS 2013). Assuming no immigration of recruits from other neighboring areas is reasonable, especially because the quality/quantity of data available for this assessment would not have allowed modeling inputs and outputs of recruits and other life stages from one subregion to another. It is also true that the stock-recruit relationship for sharks should include a lag phase, which in fact may even be longer than the gestation period if there is any form of sperm storage, to allow females which become pregnant to gestate and produce offspring. Other shark stock assessments have assumed that recruitment occurs at age 1. Expressing R_0 as density/area rather than numbers does not seem to have affected assessment results, as the author indicated, just constrained the densities to be close to each other in adjacent subregions by adding a penalty term in the corresponding likelihood.

Parameter estimation—Fixing most of the parameters was justified, as is done in similar models with similar data limitations, leaving virgin recruitment and vulnerability parameters as the only ones to be estimated by the model as recruitment deviations could not be estimated either in this case (see Recruitment section above). To estimate uncertainty, the assessment used MCMC with a thinning rate of 50, but I note that there was no burn-in phase.

Data inputs—It is unclear what the use of these CVs as slower bounds accomplished, especially considering the lack of fit of the model to the catch rate data (see Stock assessment results section 2.5). I am a bit confused by how the observer program data were used because Chapter I gave the impression that the aggregated 2006-2012 FOP data on species composition and length distributions had been used for each subregion, but the last paragraph of Section 5.7.1 seems to indicate that only the year with the most sharks observed was used in each subregion, which would mean that the information for the remaining years was not used. Or perhaps it means that the combined 2006-2012 data were assumed to correspond to the year with the most shark observations, but I fail to see the relevance of this statement since species and length compositions in each subregion were assumed to be constant. *It would have been useful to show the annual species and length compositions by subregion from the FOP as an appendix to get a sense of their interannual variability (ToR4).*

Likelihoods—1) The likelihood for the relative abundance measures used the CVs defined earlier as inverse weights, thus the more precise (smaller CV) index values were given more weight than the less precise ones. The assessment thus effectively used inverse CV weighting. I do not see a catchability term (constant of proportionality) for the vulnerable biomass in the likelihood equation 5.12. Is it assumed to be part of the vulnerability term used for vulnerable biomass (2 equations above 5.12)?

2) The likelihood for length compositions estimated an effective sample size. This effective sample size was estimated with a method that adjusted the multinomial likelihood making use of the "raggedness" of the length-frequency distributions and which appears to have been derived by the author. The Francis (2011) method was not used due to its alleged complexity. As far as I can tell the method derived by the author is sound and represents an alternative to the Francis (2011) method. The Francis method uses the mean length (age) composition residuals, as opposed to using the raw residuals as in McAllister and Ianelli (1997), from the model output and iteratively reweights the length compositions until there is no significant change in model fits and outputs. His approach also prioritizes weghting abundance data before composition data.

3) The likelihood for species compositions was treated the same way as that for length compositions.

4) The likelihood for recruitment parameters included a penalty term that constrained recruitment densities to be close to each other in neighboring subregions, which may have the effect of narrowing the confidence intervals when adding the biomass and yield estimates by subregion, as recognized in the assessment. It also affected the fit to the species composition data (see Stock assessment results section 2.5).

There was no likelihood component for the catches, which means that they were not fitted by the model and were assumed to be known without error.

Maximum Sustainable Yield—The computation of MSY used standard methods, except that the recruitment ratio (R/R_0) was expressed with a formulation of the Beverton-Holt equation that took advantage of the recruitment compensation ratio obtained in the Demographic analysis section. The fishing mortality rate that would result in the population being at MSY (F_{MSY}) was computed as the lowest value of all the species present in each subregion to follow a precautionary approach and because it was thought that it was not posssible to target a particular species in any subregion, or in other words, to fish species at different levels. While in reality F_{MSY} is an intrinsic property of each species, which is a function of gear selectivity and life history, the assumption of equal harvest rate for all species in a subregion was due to the fact that catches and catch rates were not species-specific, but aggregated, in each subregion. This assumption should result theoretically in more conservative estimates of MSY.

2.5 Stock assessment results

Overview

The model yielded a very wide range of results, predicting extremely large population sizes at the upper end of the spectrum of population biomass, presumably because the catch rates were mostly increasing or flat. This led the author to focus on the lower end of the spectrum of MSY values (the lowest 50% of total MSY estimates from the MCMC runs) because the maximum-likelihood estimate (MLE) was very large (58,330 t) and deemed unrealistic. By plotting MSY vs. the negative log-likelihood (NLL) values, two representative points were defined:

1) A substitute maximum-likelihood point (Representative Parameter Vector 1), which represents the preferred MSY estimate of ca. 5,000 t as defined by the lowest NLL values. Lower MSY values had higher NLL values and were therefore less likely, whereas higher MSY values, some of which had similar NLL values, were too far away from, and thus incompatible with, historical catches and deemed unlikely.
2) A minimum MSY point (Representative Parameter Vector 2) of ca. 1,300 t, corresponding to the low end of MSY values, but which had lower NLL values, and was therefore more likely, than points with even lower MSY values.

Model outputs (MSY and biomass estimates) were then shown only for these two scenarios, which were considered to be more plausible.

Detailed findings

Model fits—The fits shown in the assessment correspond to the substitute maximum likelihood scenario (Representative Parameter Vector 1) defined above.

1) *Fits to standardized catch rates*—The model was unable to fit the standardized catch rates for any of the 10 subregions, showing a flat tendency in all cases.
2) *Fits to species compositions*—The model fit the species composition data fairly well, with the exception of the Stanage and RockEst subregions. Since the species composition was assumed without error, the model would have fitted it exactly had the population size of species in neighboring subregions not been constrained to be similar, as recognized in the assessment.
3) *Fits to length compositions*—The assessment did not show the fits to the length compositions. Instead, the vulnerability functions fitted to each of the 12 species groups were shown (see below).
4) *Vulnerability estimates*—Fitted logistic (for small whalers) and dome-shaped (for large whalers and hammerheads) functions were plotted for the 12 species groups. The plots show that adults of small whaler species are fully selected by gillnets, whereas for larger species the greatest retention occurs at sizes well below the length

at maturity thereby protecting the adult stock. In that respect, it would have been helpful (*ToR4*) to *show the L_{50}, in addition to the length at birth (L_0) and female L_∞, in these plots to more easily visualize the proportion of the adult stock that is vulnerable to the gillnets*. The assessment indicates that three species (common blacktip, spinner, and bull) show vulnerability peaks at length at birth or very close. I also see the winghead shark and the scalloped hammerhead as falling in that category. It is not possible to judge the fit of the selectivity curves to the data because *the observed length frequencies are not plotted in the graphs (ToR4)*. So, for example, does the fit for the common blacktip mean that the data included lengths below the length at birth or that the model did not fit the data well? In either case this is worrisome because it indicates unreliability of observed lengths or lack of model fit. Based on this, my prediction is that if the fit to the length compositions had been shown, these fits would have been poor.

5) *Recruitment—The output did not show the estimated virgin recruitment (R_0)* (*ToR4*).

Model convergence—Trace plots of the 10,000 retained MCMC iterations for MSY, NLL,and the two selectivity parameters, L_{50}, and L_{diff}, for hardnose sharks, were shown and displayed no obvious trends. This was the only diagnostic shown to support convergence of the MCMC algorithm as, for example, *parameter autocorrelations were not shown (ToR4).*

Model outputs—The assessment showed MSY- and biomass-related estimates:

1) *MSY*—MSY and other biomass- and *F*-related quantities that would correspond to MSY fishing were tabulated by population and by subregion (as defined by the limiting population with the lowest F_{MSY}) and then added across species for three areas (Gulf, Northern, and Southern). Since the limiting species was great hammerhead in most cases, the corresponding F_{MSY} and MSY values for the subregions were low, but F_{MSY} was high (>0.6) for some individual populations, particularly those in which adults are not vulnerable to fishing. I find the B_{parent}/B_{virgin} ratios under MSY fishing listed counterintuitive because less productive species would be expected to have higher ratios, to the right of 0.5, whereas more productive species would display the opposite trend. For comparison, the spawning potential ratio at maximum excess recruitment (SPR_{MER}) (Brooks and Powers 2007, Brooks et al. 2010), although not exactly the same as SPR_{MSY}, predicts this trend, i.e., higher values the less productive the population. The total combined MSY for the three regions combined was ca. 4,900 t for Representative Parameter Vector 1 and 1,273 t for Representative Parameter Vector 2.
2) *Biomass*—Exploitable, current, and virgin biomass levels were tabulated in the same manner as for MSY. One notable aspect that was not explicitly mentioned in the assessment is that the values of $B_{current}/B_{virgin}$ are all very high (0.82-0.99 for Representative Parameter Vector 1; 0.44-0.96 for Representative Parameter Vector 2,

but most values>0.75), indicating that even in these two highly conservative scenarios (low assumed MSY) there was very little depletion from virgin levels.

In conclusion, considering the whole spectrum of results, the model is interpreting that the population can sustain very high catches (MSY) because of the generally flat or increasing catch rates and the more recently (since 2003) reduced catches. The problem is that the catches are likely underreported and we do not really know what the catch rates represent. Additionally, the model was essentially unable to fit the catch rate data and probably the length composition data. I also note that there were no specific reference points specified to assess an overfished or overfishing condition. In terms of biomass, only $B_{current}/B_{virgin}$ values were presented (not $B_{current}/B_{MSY}$), which showed no sign of any sotcks being overfished; in terms of fishing mortality, only F_{MSY}, but not $F_{current}$, was specified. No projections of future stock status were undertaken likely due to the uncertainty in current stock status.

2.6 Management implications

<u>Overview</u>

Based on the results of the assessment, which showed that MSY (1,077 t) was well above the current total allowable commercial catch (TACC) of 600 t for the Queensland east coast even for the most conservative scenario considered (Representative Parameter Vector 2), it was concluded that current catches (for 2013) in the region, which amounted to 157 t for the Northern Management Region and 80 t for the Southern Management Region, were sustainable. The MSY estimate for the Gulf with Representative Parameter Vector 2 (196 t) was a little lower than the reported catch of 221 t.

The assessment cautions of any further expansion of the gillnet fishery in the region because of the negative effect that potential changes in gear technology could have on the species exploited. The fishery currently targets small individuals and thus only juveniles of the larger whaler species are vulnerable to the fishing gear, which according to the assessment results seems to be a sustainable strategy. In contrast, small whalers are vulnerable during most of their lifespan by virtue of their smaller length and could thus be more at risk of overfishing. Hammerheads would be at risk of overfishing if the fishery expanded because of their higher vulnerability to the gillnets.

<u>Detailed findings</u>

With the exception of the MSY estimate for the Gulf being about 10% below the reported catch in 2013, all conclusions (*ToR2*) on the sustainability of the exploited stocks based on the assessment results are justified, especially considering that Representative Parameter Estimate 2 is a very conservative scenario. However, the results of the assessment must be viewed very cautiously owing to the data limitations and the extremely high uncertainty in estimated quantities. For that reason, the precautionary advice of being very careful with expanding the fishery should be an unambiguous management recommendation of *not allowing any increases in catch levels until more credible assessments can be undertaken*.

The view that the fishery strategy of targeting small, immature individuals is sustainable is supported by the assessment results, but those hinge on the assumption that the reported catches are correct, and more importantly, on the crucial assumption that the catch rates represent the real abundance trends of all the species occurring in the area. This caveat is recognized in the assessment, which cites changing market conditions, targeting, and reporting practices as possible reasons for the increasing trends. As concluded in this review, one main reason for the lack of confidence in the catch rate trends is species composition, which makes it impossible to interpret these trends.

2.7 General conclusions and recommendations

The assessment did a very thorough job at attempting to assess the status of a large suite of shark species occurring in Queensland waters, but was hampered by data limitations. Much of the problem stems from the difficulty in correctly identifying different species of sharks. Another serious handicap is the reliance on logbooks, which are known to be subject to different intended and unindented sources of error, and which in this case did not include discarded animals and included only partial information on fishing effort. These limitations left the Fishery Observer Program as the only reliable data source. However, the FOP also suffered from limitations, specifically its short duration and the fact that is was voluntary. Furthermore, fishing effort information was not available because it is hard to quantify for gillnet fisheries. The only more reliable source of information was thus life history. To fully make use of the biological information available for many of the species caught in the fishery, a new demographic model was developed that had the advantage of specifying stock-recruitment parameters in terms that were more amenable for use in the stock assessment model. Although the method relied on some assumptions to calculate the recruitment compensation ratios that may not be the most adequate for sharks, the effect of these assumptions can easily be investigated through sensitivity analysis, which will yield lower estimates of productivity and affect the reference points calculated, likely showing more depletion and lower MSY values. However, these changes will almost certainly not change the conclusions from the assessment given the other data limitations noted.

In all, the assessment used catches that were likely under-reported and catch rates that did not represent the relative abundance of any of the species represented in each subregion or even the species complex as a whole. The assessment considered 10 subregions and 61 populations in an effort to make full use of the available life history information and data from the FOP. While this was commendable, the CPUEs derived for each subregion were affected by changes in species composition with time and different trends for species with different productivities. As a result, the assessment model was unable to fit the CPUE series, and likely the length compositions, and found that there had hardly been any depletion from unexploited levels and a huge range of possible values of MSY. By taking a risk-averse approach and focusing on the most likely and lower values of predicted MSY, the assessment still found that current catches and catch limits are sustainable.

In terms of the desired objectives listed for the stock assessment, they included $B_{current}/B_{virgin}$ as a biomass reference point and F_{MSY} as a fishing mortality rate benchmark for all stocks considered, but the $F_{current}/F_{MSY}$ and $B_{current}/B_{MSY}$ reference points were not included. While Total Allowable Catch (TAC) was not presented, MSY was generated for each stock and compared to the current commercial TAC to show that both current catch (for 2013) and TACCs were below MSY and thus sustainable based on the results of the assessment. In all, the assessment fulfilled its objectives to the extent possible given the data limitations.

The assessment concluded by enumerating a number of potential improvements to both the analytical methods used and the input data available. While most of the improvements listed could be beneficial to different degrees, I consider them to be only of second-order importance. The bigger issue is the quantity and quality of the fishery data available as also recognized in the assessment. With the current data limitations, future assessments of Queensland shark resources will continue to be extremely uncertain and not be able to provide very effective management advice. While there are a number of practical recommendations about the data already available and model used that could be useful, such as accounting for dead discards and post-release mortality in gillnets using the existing observer data or considering a Bayesian approach to parameter estimation, or other data or modeling suggestions that have been presented earlier in this review, I would like to focus on **two main recommendations** that would allow for a significant improvement of the assessment. First, there has to be a serious investment in data collection. In this reviewer’s opinion the most effective way to collect the additional information required to conduct an improved and more credible assessment of shark resources in the area is to reinstitute the ***Fishery Observer Program***.This would provide crucial pieces of information on the annual catch of gillnet fisheries, including species compositions, length compositions, status and fate of captured animals, development of an alternative relative abundance index, and collection of biological samples for life history and genetic studies. Alternatively, a fishery-independent survey of shark resources could be initiated covering the entire area and with similar objectives to those just mentioned, but this would likely require collaboration from academic partners and standardization of methodology. Second, since distribution of field guides (for whole animals or carcasses) is not thought to be an effective strategy for species identification by fishers or dock-side personnel, another relatively low-cost strategy could be to take ***fin clips*** from specimens captured in the gillnet fishery to genetically identify them and then compare the species ID to those reported in the logbooks. This would allow verification of the species composition in the logbook data, attempts at reconstructing the species composition back in time, and adjustment of newly acquired logbook data.

3. References

Au, D.W., Smith, S.E., and C. Show. 2008. Shark productivity and reproductive protection, and a comparison with teleosts. Sharks of the Open Ocean (eds M.D. Camhi, E.K. Pikitch & E.A. Babcock), pp 298-308. Blackwell Publishing, Oxford.

Brooks E.N., and J.E. Powers. 2007. Generalized compensation in stock–recruit functions: properties and implications for management. ICES J. Mar. Sci. 64, 413–424.

Brooks, E.N., Powers, J.E., and E. Cortés, E., 2010. Analytical reference points for age-structured models: application to data-poor fisheries. ICES J. Mar. Sci. 67, 165–175.

Cortés, E. 2000. Life history patterns and correlations in sharks. Rev. Fish. Sci. 8:299-344.

Courtney, W.D. 2013. A preliminary review of post-release live-discard mortality rate estimates in sharks for use in SEDAR 34. SEDAR 34-WP-08. Southeast Data, Assessment and Review (SEDAR) 34: stock assessment report—HMS Atlantic sharpnose shark. North Charleston, SC, USA.

Francis R.I.C.C. 2011. Data weighting in statistical fisheries stock assessment models. Can. J. Fish. Aquat. Sci. 68: 1124–1138.

Goodyear,C.P. 1977. Assessing the impact of power plant mortality on the compensatory reserve of fish populations. In: van Winkle W (ed.) Proceedings of the Conference on Assessing the Effects of Power Plant Induced Mortality on Fish Populations. Gatlinburg, Tennessee, 3–6 May 1977. Pergamon Press, New York, pp 186–195.

Goodyear, C.P. 1980. Compensation in fish populations. In: Hocutt, C.H., Stauffer J.R. (Eds.) Biological monitoring of fish. Lexington Books, DC Heath and Co, Lexington, Massachusetts, pp. 253–280.

Gulland, J.A. 1970. The Fish Resources of the Ocean. FAO Fisheries Technical Paper No. 97 FAO, Rome.

Lo, N.C., Jacobson, L.D., and J.L. Squire. 1992. Indices of relative abundance from fish spotter data based on delta-lognormal models. Can. J. Fish. Aquat. Sci. 49:2515-2526.

Maunder, M.N., and A.E. Punt. 2004. Standardizing catch and effort data: a review of recent approaches. Fish. Res. 70:141-159.

McAllister, M.K., and J.N. Ianelli. 1997. Bayesian stock assessment using catch-age data and the sampling-importance resampling algorithm. Can. J. Fish. Aquat. Sci. 54: 284‑300.

Myers, R.A., Mertz, G., and P.S. Fowlow. 1997. Maximum population growth rates and recovery times for Atlantic cod, *Gadus morhua*. Fish. Bull. 95, 762–772.

NMFS (National Marine Fisheries Service). 2002. Stock assessment of large coastal sharks in the U.S. Atlantic and Gulf of Mexico. NOAA/NMFS/Panama City Laboratory, FL, USA.

NMFS (National Marine Fisheries Service). 2013. Southeast Data, Assessment and Review (SEDAR) 34: stock assessment report—HMS Atlantic sharpnose shark. North Charleston, SC, USA.

Pardo, S.A., A.B. Cooper, and N.K. Dulvy. 2013. Avoiding fishy growth curves. Methods Ecol. Evol. 4:353-360.

Sainsbury, K. 2008. Best practice reference points for Australian fisheries. Report to the Australian Fisheries Management Authortity. December 2008.

Smith, S.E., Au, D.W., and C. Show. 1998. Intrinsic rebound potentials of 26 species of Pacific sharks. Mar. Freshwater Res. 49, 663–678.

Zhou, S., Yin, S., Thorson J.T., Smith, A.D.M., and M.Fuller. 2012. Linking fishing mortality reference points to life history traits: an empirical study. Can. J. Fish. Aquat. Sci. 69, 1292–1301.

Appendix 1: Terms of Reference

Conduct a review of the "Stock assessment of whaler and hammerhead sharks in Queensland". The review is not limited, but should address, the following points:

1. Review the report inputs and outputs and adequacy of these data in order to achieve the objectives of the assessment, including:

   (a) The new demographic analysis and the demographic parameter values input to the population model;
   (b) Use of data from the Fishery Observer Program which was the primary data source of data for species identification, sizes of sharks caught and discard rates for the commercial catch (this program employed government observers to record data on commercial fisheries from 2006 to 2012);
   (c) Non-use of species identification records from commercial logbooks;
   (d) Estimated historical harvest levels going back to the mid-1970s;
   (e) The generalized linear model and explanatory variables used in the commercial catch-rate standardization to derive annual abundance estimates;
   (f) The adequacy of the population dynamic model used in the assessment;
   (g) The model's estimates of different species' vulnerability to fishing and implications for susceptibility of the populations to fishing pressure;
   (h) Any further caveats or assumptions that may need to be considered.

2. Comment on the accuracy of all the statements made in the Stock Assessment, but particularly the section "Imlications for fishery management" (section 7.1 in the Discussion chapter) and how well they are supported by available data and model results.

3. Potential sensitivity analyses that could have been conducted.

4. Any other outputs or graphical figures that the report could have provided.

A formal written report of the findings of the review is to be provided to the stock assessmennt author and a nominated person from Fisheries Queensland. The written review and review author identification may be released and made publically available.