# Response to reviewers’ comments

We have no reason to doubt that the changes to the conditioning of the operating models were important for improving fit to the data; however, the effects of many of these updates are hard to judge given the available documentation. For example, how does the new trawl selectivity function (Gamma shaped) compare to the previous selectivity function? How much of an impact did this have on the operating model conditioning? How much of an impact did the new assumption about tighter observation errors on the trawl at-sea release have on the operating model? What do the inferred trawl age-composition data look like for selection of years and what do the length-composition data look like that go into that? What do the raw aging error plots look like (Reader 1 vs. Reader 2)? This might be more than typically goes into a Science Response, but there are quite a few changes made to the operating model here and we imagine some documentation bridging the operating model changes will be helpful for future analyses.

We will include a comparison of the two trawl selectivity function parameterisations, and some plots of the inferred trawl age compositions. The selectivity functions are not that different, given that the fully selected age class is high enough to reduce the effect of the gamma density’s skew, but we found that the ADMB operating model was unable to converge using the normal density kernel, whereas the gamma density function was more well behaved in the optimisation. An exhaustive bridging analysis is impossible given the time constraints here.

Given the importance of the very high 2015 estimated year class throughout this document, we would like to see some diagnostic plots to help understand what led the model to estimate this. The Response notes high trawl at-sea release observations. To what extent is this evident in age composition data?

StRS age compositions from 2017/2018 show a sizeable age 2/3 classes, as do the inferred trawl age compositions from the age-length key. However, it is still too early to see a large effect in the offshore commercial trap samples.

As an aside, there may be some qualitative confirmation in the synoptic trawl survey data, where the surveys are largely catching juvenile sablefish. All four surveys show some level of increase in biomass density after 2015 (although maybe not Hecate Straight in 2019) and the length distributions are available. Queen Charlotte Sound and West Coast Vancouver Island are the clearest.

Thanks for the info. This helps support our choice of the high recruitment scenario as the reference OM set, as opposed to the more conservative low recruitment scenario

As noted in the results, one important reason that the retention policies examined in the document are able to meet conservation objectives while sustaining higher catches and higher fishing mortalities is that they reduce growth overfishing contingent on fishing stopping once the TAC is met. The end of the methods notes "The most critical assumption, and even unrealistic in some cases, in the above is that fleets stop fishing when their fleet-specific TACs are fully landed." Is there currently no implementation error? If so, this seems like an important dimension to explore in a future update.

This line is mostly in reference to the multispecies nature of the trawl fishery. It is unlikely that they will stop fishing when their juvenile TAC is met, and it is unclear if management actions to reduce TAC to zero from juvenile release overages would in fact stop the trawl fleet from fishing in any given year. It is likely that we should explore this avenue further, however we predict it is probably very similar to the status quo, if the simulated incentives are essentially toothless. As for general implementation error, we think this is unnecessary for BC groundfish, and especially so for directed Sablefish fishing. Any systematic implementation error for the legal sized TAC is unnecessary, given the the at-sea observer programme and electronic monitoring combined with 10% logbook auditing by Archipelago. Any random error would probably be indistinguishable from the median behaviour.

The section in the results on cross testing that is yet to be written seems particularly important. The summary in the conclusions and the tables themselves make it clear that there is a great asymmetry in risk here. The consequences of overestimating a large 2015 year class has real conservation risks.

We agree and recommend the risk-averse approach of harvest rates tuned to meet objectives under the low recruitment scenario.

We assume there is a historical reason for dividing the steepness-SSB joint posterior into 5 regions and then recombining them afterwards. In this current document, the 5 operating models are never considered independently and always as a weighted average. Separating them and then combining them weighted by their marginal density would seem to only add complexity, add some error from the arbitrariness to the divisions, and—perhaps most importantly—not sample from the outer tails of the posterior. The only advantage we can think of is that it could make communication easier by being able to consider the components individually. Presumably other longer documents in the future will take advantage of the separation. Otherwise, it would make sense to me to just sample from the full joint posterior.

This is accurate, in the last CSAS paper we ran 5 OMs based on a single draw that was from the posterior closest to the defined centres (in a Euclidean sense), rather than sampling 100 points around those centres. We extended this to a sampling procedure based on reviewer comments from that paper but have since learned that random sampling and increased sample size is probably superior for adequately capturing risk, rather than attempting to stratify in this way. Although your concern about missing the tails is founded for the pair of parameters defining the joint marginal that we stratified, we found that the posterior means and standard deviations for most other model parameters and deviations in the 5 strata were similar to those from the full posterior (compare Mm and Mf across 5 OMs in Table 1). We will likely reduce complexity by sampling 500 points from the full posterior in future.

A main conclusion seems to be that the current management procedure with no cap on at-sea releases is able to meet conservation objectives but that management procedures that involve some retention policy could also meet those conservation objectives while increasing catch and fishery revenue (currently written as income). The economic conclusions are contingent on the price structure for sablefish size and also on constant costs to the fishers (if income). Is there reason to think that a retention policy would affect average trip costs? Presumably there is considerable uncertainty about future price structure for sablefish size. This may be worth a brief discussion.

You are quite right, income should be revenue. We decided that the research scope on this project was too limited to incorporate variable costs of fishing for this calculation. We agree that different retention policies would also likely affect trip costs, either through fuel (avoidance behaviour or move-on policies) or sorting time (increased recording and retention regulations).