Center for Independent Experts (CIE) Individual Review Report for the review of Eastern Bering Sea Tanner Crab Assessment, Seattle, USA, 9–11 June, 2026

Martin Cryer, August 2026

Contents

Executive Summary

This is a review of aspects of the 2025 Tanner crab assessment and how it might be developed and used in the future for developing management advice. For ToR1, I think all future models should be developed using the GMACS framework. For ToR2, although I believe the methods used to estimate the catchability and selectivity of the NMFS survey gear are broadly sound, and the resulting changes in population scale in the assessment model are defensible, additional work could be done to bolster those estimates. Such work could include some or all of exploring trawl "lift-off lags", pooling data for all *Chionoecetes* crabs, doing depletion experiments, fitting models direct to the data instead of to ratios, and using curves more typically used in stock assessment models rather than heuristic models. For ToR3, I don't believe that the current $F_{35\%SPR}$ reference point is useful or defensible for this combination of biology and fishery, and I believe a more empirical, simulation-based approach would be better and a high priority. If reference points are to be reviewed, I think an assessment of what led to population crashes for some "stocks" in Alaska while others persisted should be explored as part of the background work. For ToR4, I think developing a multi-area model for EBS Tanner crab is a low priority. Survey data can be used directly to inform management at finer spatial scales if this is thought desirable. For ToR5, I don't think the recent high prevalence of hybrids is a cause for much concern, although monitoring should continue. No sustainability risk is generated by the current practice of excluding hybrids from the Tanner crab assessment and I believe additional catch targeted on hybrids could be provided for in other ways. Finally, the current stock assessment model has a striking (and intriguing) inability to reproduce very believable peaks in the survey biomass index and the reasons for this should be investigated. I suggest exploring data weighting, survey catchability issues, potential environmental drivers, and density dependence.

Background

Tanner crab, *Chionoecetes bairdi*, are found in continental shelf waters of the north Pacific, including, at the northern limit of their range, the Eastern Bering Sea (EBS) where water temperature may limit the northern extent of their spatial distribution. The distributions of Tanner and snow crab, *C. opilio*, overlap from ~56° to 60°N where the two species hybridize. Both Tanner and snow crab show marked variations in their abundance, driven by variable recruitment, and fishery closures are common.

Currently (in 2026), snow crab abundance is low whereas Tanner crab abundance is increasing strongly. Also, historically high proportions of hybrids of the two congeners have been observed in surveys. If these trends continue, Tanner crabs and hybrids seem likely to become more important components of the crab fisheries in the EBS than hitherto. This review was designed to help ensure the science used in the management of Tanner crab continues to be robust and transparent in the pursuit of management advice that maximizes long-term sustainable catches.

Role in the Review

Before the meeting, I read the documents and viewed the presentations provided on a github repository. I also retrieved and studied the 2023 and 2025 stock assessment SAFE documents for this Tanner crab stock (after the review meeting, I also looked at the 2000 SAFE document which was very useful).

During the review meeting, all panel members participated in the discussions and the chair was diligent in ensuring each panel member's views were heard for each ToR. In a sense, this was an unusual review because the ToR did not include a formal review of the stock assessment model itself but, rather, covered future development of the model and translation of modeling results into management advice. These matters are not always easy to disentangle from the modeling itself, and we occasionally found ourselves discussing particular aspects of the model fits and diagnostics insofar as they related to the other ToR. This was particularly the case when considering the side-by-side selectivity and catchability information (ToR2) and how to include it in the assessment model.

My background and experience in crustacean and bivalve stock assessment, including estimating survey selectivity and catchability, led to me having a bit more to say on ToRs 1–3 than on ToRs 4 and 5, but I am confident that, collectively, the panel had a sufficient span of experience and expertise to cover all ToRs in this review.

I enjoyed the review meeting and found both the review team and the stock assessment team to be open, very thoughtful and respectful. The stock assessment team are clearly very skillful and thorough in their explorations and at no time did I see any hint of defensiveness. It was a genuine pleasure.

Summary of Findings by ToR

ToR1: Generalized Model for Assessing Crustaceans Stocks (GMACS)

This ToR called for an assessment of the pros and cons of transitioning from the existing bespoke model to the GMACS framework and whether there are any concerns with the latter framework that would preclude its adoption.

Without in any way wanting to denigrate the current bespoke model TCSAM02 and the current stock assessment, it seems to me that shifting this assessment to the GMACS framework would have many distinct advantages and few, if any, disadvantages. Both use the ADMB optimizer which is now obsolescent and will soon not be supported. However, I understand that GMACS will be transitioned to the RTMB optimizer and it makes sense to do that work for a single package rather than for multiple bespoke models. Similarly, software maintenance would be simplified if all crab stock assessments were conducted using the same package, reducing the likelihood of bugs or inconsistencies between assessments. I understand the GMACS package already includes options for multiple fisheries and spatial structure so, if these were considered potentially useful for Tanner crab assessment, they could be much more easily explored in GMACS than in TCSAM02. Moving to a GMACS framework, therefore, would provide for greater simplicity, consistency, transparency, efficiency, and accessibility for analysts.

Differences in the way TCSAM02 and GMACS treated some parameters and calculations meant that it has taken some time, and a significant investment by the Tanner crab assessment lead scientist into understanding precisely what the code in GMACS does, to achieve the same results with the two models. However, those obstacles have been overcome and work started on fitting in GMACS during this review. The first run did not successfully converge but such teething problems are not unusual and should not be hard to overcome. Overall, I believe that the disadvantages of shifting to GMACS that did exist have been surmounted by adding features and options to the GMACS framework such that TCSAM02 no longer has any advantages over GMACS. No further effort should be expended on the development of TCSAM02.

In summary, I believe the pros of moving to the GMACS framework now greatly outweigh the cons and I see no concerns that would preclude using GMACS.

ToR2: Including side-by-side selectivity and catchability information in the model

This ToR called for an assessment of whether the methods for estimating the (catchability and) selectivity of the NMFS survey gear based on the side-by-side experiments are scientifically sound and whether the resulting changes in population scale of the assessment model are defensible and appropriate for providing management advice.

General description

The lead assessment scientist presented a very detailed analysis of the data and catchability / selectivity analyses for the NMFS standard survey bottom trawl gear (an 83-112 otter trawl) used for Tanner crab. That gear is widely considered to be size-selective for Tanner crab *via* escapement under the footrope but the BSFRF Nephrops trawl fished alongside at each station in the study has better bottom-tending abilities and is generally assumed to be non-size selective for Tanner crab across the range of sizes of interest to the stock assessment (25–200 mm carapace width). Paired tows were conducted each year between 2013 and 2018 (35–102 sites in a year, collectively covering a large proportion of the total survey area) and a total of 384 pairs were completed. All crab were counted, measured, and allocated to appropriate sex and maturity groups other than for very large samples (>~ 300) which were subsampled and re-scaled appropriately. The effective sampled area for each tow was estimated using bottom contact sensors to determine distance on the bottom and net sensors to determine net width while fishing. This represents a very powerful data set for estimating catchability and selectivity of the NMFS gear relative to the Nephrops trawl. Additional assumptions permit the estimation of absolute catchability and selectivity. Survey-level (aggregated) and haul-level analyses were presented using a variety of models fitted to the ratio of catches in the two gears.

Haul-based approach generally preferred

Instinctively, given the impressively large number of paired tows, my preference would be for a haul-based analysis designed to estimate catchability by length, sex, maturity, and in relation to environmental and between-year factors that could potentially affect catchability (depth, temperature, sediment type, etc.). This is not to say that selectivity and catchability would necessarily be implemented in the assessment model in such a complex manner, rather that the best possible understanding of the influential factors could be achieved (and corrected for, if necessary). The survey-based approach avoids some of the computational disadvantages of such an approach (see below), but it provides few replicates at each combination and, because the SBS surveys covered different grounds each year, it seems difficult to me to assess any between-year effects and their drivers. This seems a big price to pay.

Unfortunately, however, because the approach used to analyse the haul-by-haul data involved fitting to ratios, a high proportion of records (the ratio of the numbers in a given sex-maturity-length bin for the two gears) could not be computed or were considered unreliable and were deleted. This is not necessarily a bad thing, and I have done it myself in the past (see below), but it can be considered arbitrary and is potentially biased. In total, 5275 records were dropped because of low sample size and only 2430 were used to fit the models. This seemed a very high proportion to me, and plots in the report seem to suggest that the dropped records were not a random selection with respect to catchability at length, especially for males (Figure 1). This may introduce bias into the estimated catchabilities, although I would expect the effect to be relatively small.

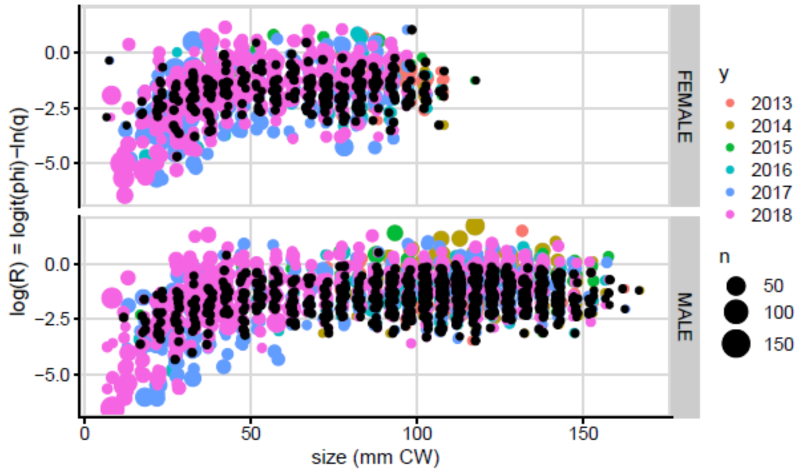

Figure 1: A copy of Figure 11 from the SBS analysis report. Haul-level ln(R) values; colours indicate SBS study years; symbol size indicates number of crabs sampled. Black symbols indicate values with fewer individuals in a size bin compared with a selection criterion. Upper plot: females; lower plot: males.

In New Zealand, my initial attempts to estimate dredge efficiency and selectivity for scallop surveys involved 32 direct comparisons of catches by dredge and SCUBA divers at "nearby sites" in the 1990's and early 2000s (e.g., Cryer & Parkinson 2006). In that analysis, I censored the data to exclude records with low sample size, calculated ratios and bootstrapped the data to estimate variances, much like the analyses presented here. I did not proceed to model the resulting curves but, rather, included the bootstrapped results empirically when correcting the survey results for catchability. A review by a technical working group concluded that those methods required arbitrary assumptions for the omission of data where samples were small and produced curves with unrealistic spikes and troughs owing to random variation in the samples. The review considered my empirical curves less appropriate for inclusion in a stock assessment than smooth, parametric curves. In response, (Bian et al. 2012) developed a complex modelling approach fitting the count data directly and allowing dome-shaped selectivity and multiple covariates such as depth and sediment type. This approach is not without its problems (see below), but it can use all the data and the impressively large paired tow data set for Tanner crab may allow such models to be successfully developed and fitted.

A detailed independent review by Smith et al. (2016) found that, although they addressed the significant issues with my initial "heuristic" analysis, the complex models developed by Bian et al. (2012) had a number of problems related to the inclusion of covariates and the very strong influence of penalties and priors (especially those related to dome-shaped selectivity). As a consequence, simpler modelling approaches are currently being used (e.g., Tuck et al. 2018) to correct scallop survey estimates in New Zealand using the SELECT method (Millar 1992) via the *trawlfunctions*[1] R-package developed by Russell Millar. Smith et al. (2016) recommended that experimental and modelling work should continue but noted that the general approach of estimating selectivity and catchability simultaneously could be problematic; modelling these two processes separately might be better since, after selectivity has been estimated,

[1] https://www.stat.auckland.ac.nz/~millar/selectware/code.html

efficiency can be obtained by estimating a single parameter (see also Tremblay et al. (2006) on the lobster *Homarus americanus* and Yochum et al. (2008) on the scallop *Placopecten magellanicus*).

I provided the key reports on this New Zealand experience to the panel and assessment team after the review meeting. I offer them not necessarily as representing better methods than those applied by the team, but as another way of thinking about things that may allow them to use more of their hard-won data. I think it is possible that, given the amount of data available and the apparent surprise of some stakeholders at the estimates of catchability, a direct modelling approach to captures at length could be explored to estimate the catchability and/or selectivity for Tanner crabs.

Other potential approaches for selectivity

The smooth GAM and other models fitted to the SBS data did a good job of broadly describing the nature of the selectivity of the NMFS gear relative to the Nephrops gear, which was assumed to be non-selective. However, I think it would be better to fit more traditional selectivity curves because this would lead to a more natural approach to implementing selectivity in the stock assessment model. Asymptotic models such as the logistic (symmetrical) or Richards (asymmetrical) selectivity curves are frequently used, although normal and double normal curves could also be considered. Based on my experience, I don't see strong evidence for domed selectivity (i.e., declining catchability with increasing size beyond some point) and I also find it hard to imagine a plausible mechanism whereby larger crabs would be less catchable for survey gear (other than, perhaps, the propensity to bury themselves increasing with size). Conversely, I agree it's better to estimate selectivity and catchability for male and female crabs separately. Male Tanner crab may be caught more efficiently than females because they have longer legs, in relation to their carapace width, than do mature females, and they are less sedentary than females and less prone to bury themselves (Somerton & Otto 1999).

The selectivity and catchability of the NMFS survey trawl gear have previously been estimated for snow crabs (Somerton et al. 2010?). I broached during the meeting what I thought was a "left-field" idea to combine data for both species of *Chionoecetes* crabs to increase the sample sizes and provide for more haul replicates to assess covariates. My rationale was that the two species appeared morphologically very similar to me, although it was discussed at the meeting that one or other species might be more likely to be soft at the time of the survey and this could plausibly affect catchability. I did not have access to the data or fitted curves, so I have overlaid some plots from the various reports as best I can by eye in Figure 2. The estimated curves for both males and females are somewhat different, but none of the modelled shapes is particularly convincing for me and the confidence limits often overlap suggesting (to me) that combining the data for some exploratory analyses might be fruitful.

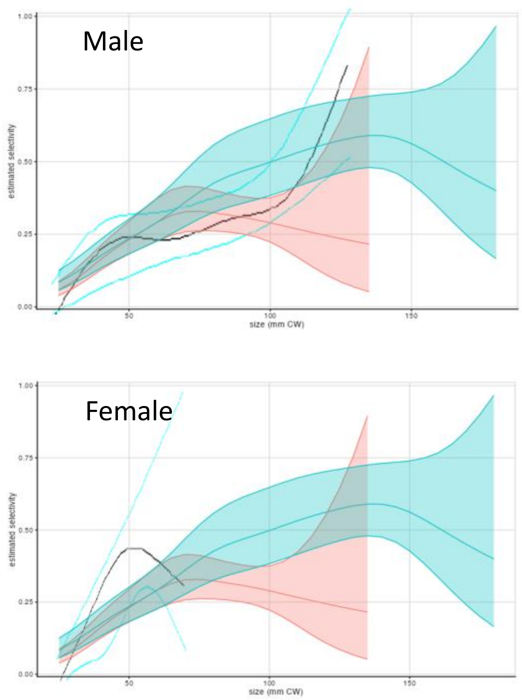

Figure 2: Best survey-level selectivity curves from Stockhausens's (2026) SBS analysis for Tanner crabs (blue = males, pink = females) overlaid (by eye) with the empirical selectivity curves from Somerton et al. (2010) for snow crab (black lines with pale blue confidence bounds) in the EBS.

Thinking further about this idea, I studied previous estimates of selectivity for *Chionoecetes* crabs in the EBS by Somerton & Otto (1999) and Somerton et al. (2013). The 1999 study, conducted using a second underbag and tickler chain below the groundrope of the NMFS survey gear, suggested fairly typical asymptotic selectivity and a fully-recruited catchability of ~0.8 for both snow and Tanner crabs (Figure 3), quite different from the latest SBS analysis presented to the review panel. The 2013 study concluded that the differences between the results of the underbag study and those from their comparative (paired tow) experiments could potentially be explained by the differences in the trawling procedures in that experiment compared with those used on the survey, plus the restricted spatial location habitat type of the former study. The differences are, indeed, large (Figure 4), and both are very different from the most recent SBS analysis for crabs of the same genus. This variability of fitted curves reflects my own experience with scallops (see also Delargy et al. 2023 for a review) and scampi (*Metanephrops*) and highlights the difficulty of estimating the selectivity and catchability characteristics of towed gear, even when lots of good data are available. For me, based on the sensitivity to modelling choices and the data scatter, it's entirely plausible that the many different estimated curves for *Chionoecetes* crabs are all just different views through quite murky windows rather than examples of real differences among species, years, or habitats.

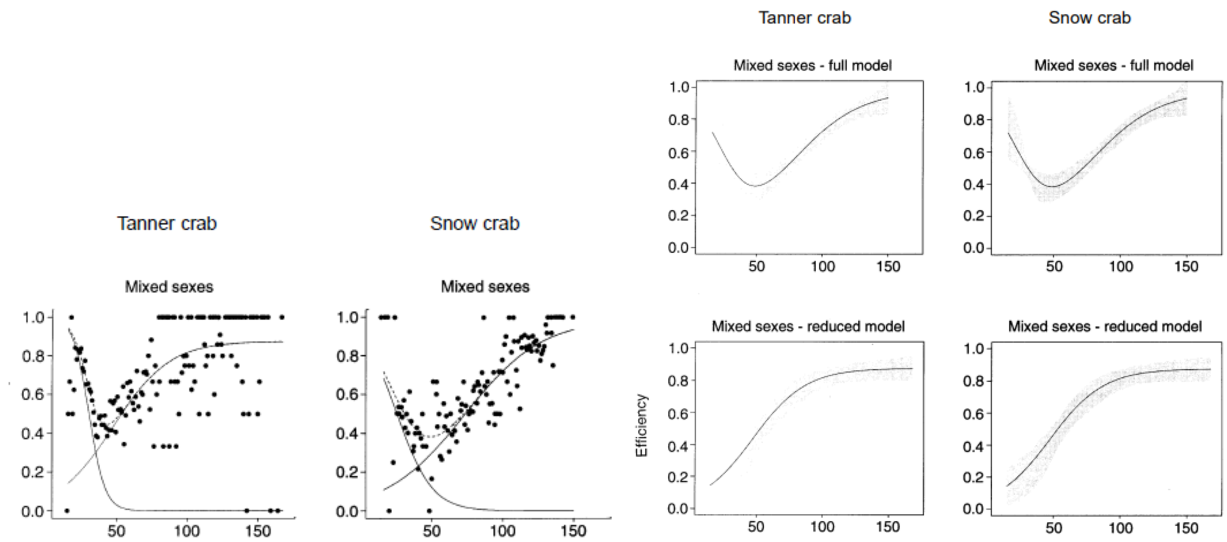

Figure 3: Estimated selectivity curves (mixed sexes) for *Chionoecetes* crabs in the EBS from Somerton & Otto (1999, their Figures 3 and 4). Left panels show the fit to the data. Right panels show the estimated selectivity curves with the top curves showing the selectivity of the main gear and the underbag combined and the bottom curves showing only the selectivity of the main gear.

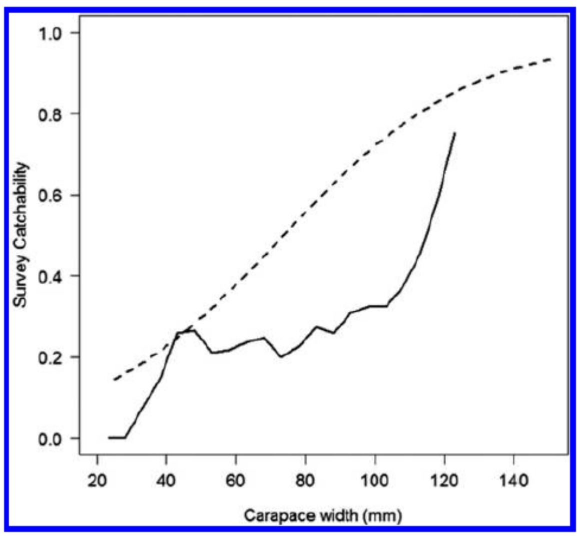

Figure 4: (copy of Figure 11 from Somerton et al. 2013) comparison of their male survey catchability function (solid line) and the male trawl efficiency function estimated in Somerton and Otto (1999) (dashed line).

Other potential approaches for catchability

Zacher et al. (2026) noted that "*Data on bottom contact of the footrope were combined with GPS data to calculate distance fished (while in contact with the seafloor), which was then combined with the net width data to calculate area swept*". I understand from discussions at the meeting that "GPS data" referred to positions acquired for the survey vessel, not specifically for the trawl gear. It occurred to me, based on my experience with Nephrops trawls in fairly deep water (200–600 m depth) that the actual effective trawled area might be larger if the trawl did not lift off the seabed very soon after haulback was initiated in these surveys. West & Wallace (2000) showed, using both simple trigonometry and acoustic trawl positioning systems (SIMRAD ITI, see also Wallace & West 2006, Engås et al. 2000) that such “liftoff lag” could affect effective sampled area quite substantially (see also Wallace & West 2006). Their study was in much deeper water than the EBS survey (200–1200 m compared with 20–200 m), but the estimated distance traversed during the early stages of haulback at a depth of 200 m was of the order of 100–400 m along the seabed at speeds similar to the nominal survey towing speed (Figure 5). They found that a simple trigonometric approach gave very similar results to using sophisticated acoustic monitoring equipment.

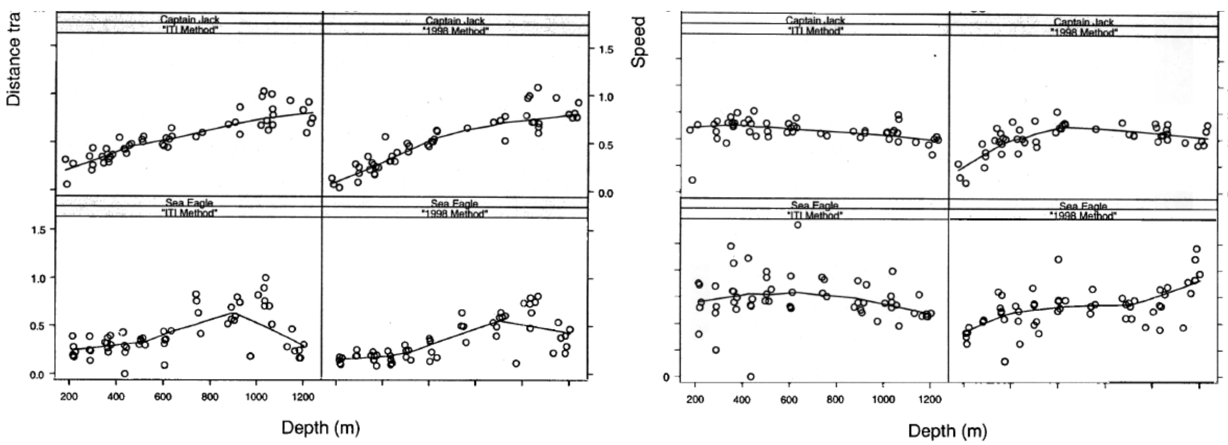

Figure 5: Distance traversed (left four) and speed along the seabed (right four) for trawls towed by two of the four vessels in the study by West & Wallace (2000) (their Figures 4 and 6).

I believe this is relevant to the SBS study because the tow durations for the two sampling trawls were very different and liftoff lag would, *a priori*, be expected to be a proportionately bigger bias for shorter tows (all other operational details being the same, which may not hold, of course). The NMFS survey gear is, nominally, towed at 3 knots for 30 minutes for a distance of 2.78 km, whereas the Nephrops gear in the SBS study was towed for only a nominal 5 minutes (for a distance of 0.46 km if the towing speeds were the same). If the average liftoff lag was, say, 200 m for both gears (I think this would be for the deepest stations only), this would lead to a bias in area swept of ~7% for the NMFS gear but a bias of 43% for the Nephrops gear. Fulanda & Ohtomi (2011) also estimated large biases for trawl at similar depths to the EBS, larger biases of up to 130% for 10-minute than for 20-minute preset tows. I don’t believe any such bias in area swept would affect the estimated shape of the selectivity curves in the SBS study unless the effective towing speed or bottom tending ability changed during the liftoff lag. However, I think it could substantially affect the catchability estimates and analysis of the effect of depth. I calculated, roughly and assuming all stations could be affected, that a substantial liftoff lag of 100–200 m, similar for the two gears, would increase the estimate of catchability for the NMFS gear from ~0.50 to ~0.60–0.67. So, contrary to my view expressed at the meeting, additional analysis of the SBS study data, and possibly some additional assessment of liftoff lags, could increase the estimate of catchability for the NMFS gear and the change may not be small.

A key assumption of the SBS analyses is that the catchability of the Nephrops trawl is known and is 1.0 for the size classes of interest to the model. Depletion experiments and/or visual methods could be used to test this proposition, or to estimate catchability at length for the NMFS survey gear directly. Given the complications introduced by burying behaviour in these crabs, and consequent difficulties observing them reliably in images or video, depletion experiments may be more feasible and can be conducted in just a few days (for one or a few substrate types) with little special equipment. Vølstad et al. (2000) estimated the efficiency of survey dredges used to catch crabs in Chesapeake Bay using two different simple depletion models. They estimated average efficiency to be only about 15% with considerable uncertainty and some differences between the two models. Rago et al. (2006) developed a more realistic spatial approach called a patch model whereby the heterogeneity of (cumulative) fishing effort within an experimental area is explicitly considered in the estimation. They demonstrated their model using data for Atlantic surfclams (*Spisula solidissima*), but it is not restricted to clams and can be applied to other sedentary species or species with relatively low motility over the duration of experimental fishing in an area (see also Hennen et al. 2012, Poussard et al. 2021, Munroe et al. 2025 for more recent developments with this approach). Wilberg et al. (2013) used a simulation approach to show that both types of depletion experiments / models work well if the location of the sampling gear is known accurately. For these models to perform at their best, experimental areas that are large relative to the expected motility of the crabs should be used, the location of the gear (not the vessel) during fishing should be known accurately (probably using acoustic technologies), and care should be taken to pinpoint the start and end of effective fishing, especially in deeper water. As I noted at the meeting, visual and depletion experiments can only be expected to lead to estimates of catchability <1 for the Nephrops trawl, meaning that this type of work can only decrease the estimates of absolute catchability for the NMFS survey trawl and further increase the population scale in the stock assessment model. This would not surprise me, given the range of catchability estimates for invertebrates I have seen over the years, but it seems like it would be difficult for some stakeholders to accept.

Including SBS information in stock assessment models

Currently, the stock assessment model includes the SBS data as two different survey indices (with compositions) and what I found to be an unusual "availability" vector imposed on the model to allow for the differences in coverage of the SBS experiments between years. I think this has worked reasonably well, but I think it would be more conventional to use the SBS data and analyses to form priors for survey q and selectivity curves. I see from previous assessment reports (e.g., Stockhausen 2020 and 2023) that this idea is not new and a variety of asymptotic and empirical selectivity curves have been tried in the past. A prior on a scalar for fully-selected catchability seems straightforward to me, but the number and type of priors on selectivity curves would depend on the shape / statistical form of selectivity curves chosen. Some thought might be required about whether using the length compositions from the two gears in the SBS study could mean some data might be inappropriately used twice (in the development of priors then in the model fitting) and, if so, whether this could be mitigated by data weighting. Most of these choices would naturally follow the adoption or agreement on the "best" curves from the SBS analyses, whether or not those include any or all of my suggested additional work.

The treatment of selectivity and/or catchability in the stock assessment model and the choice of priors could become particularly challenging if selectivity and/or catchability is found to be a complex function of variables like depth, bottom temperature, sediment type or they vary between years for unknown reasons. I'm not sure whether it would be worthwhile to attempt implementing this complexity in a stock assessment model, or even if that would be possible in GMACS, and some simplification would probably be better once the implications have been considered.

Summary of thoughts on SBS analyses

In summary, I believe the methods used to estimate the catchability and selectivity of the NMFS survey gear based on the side-by-side experiments are broadly sound and I believe that, noting some ongoing uncertainty, the information has been included in the stock assessment model appropriately. Although I have some reservations about the area swept assumptions in the SBS experiments, there is a strong and consistent signal on the different performance of NMFS vs Nephrops gear in the SBS data which means that quite large changes in population scale in the assessment model are to be expected and defensible. The model is, therefore, suitable for providing management advice, noting that the current mechanism for calculating such advice (see ToR3) does not seem appropriate. I believe additional work could be done to bolster the estimates of catchability and selectivity of the survey gear, potentially making those estimates more believable for stakeholders. Contrary to my view expressed at the meeting, further reading and a few simple calculations has led me to believe that such work may increase or decrease the NMFS survey catchability estimates and lead to smaller or larger changes in population scale than exhibited by the current stock assessment model.

ToR3: Reference points and harvest control rules

This ToR called for an assessment of whether $F_{35\%}$-based reference points are appropriate for Tanner crabs, advice on potential alternative reference points, and the choice of biomass metrics for use in harvest control rules.

Background to the current reference points

$F_{35\%}$ was developed for groundfish (Clark 1991, 1993) to permit high yield across both sexes and a relatively wide spread of age/size classes with similar maturity and fishery selectivity ogives. Both domed and asymptotic S-R relationships were considered in the simulations but their shape was constrained. Subsequent work (e.g., Clark 2002, Saud 2007) suggests that $F_{35\%}$ is very sensitive to the assumptions about maturity and selection. Clark (2002) noted that his initial paper (Clark 1991) did not consider effects on stock biomass at all, although his later paper did (Clark 1993) and "*reported that biomass levels below 20% of the unfished level would not be uncommon in stocks with strong serial correlation in recruitment that were fished at $F_{40\%}$*". Thus, a $F_{35\%}$ constant-F strategy seems likely to lead to low biomass of (mature male) crabs at times and, as Clark (2002) pointed out, maintaining biomass above B_{MSY} (or a proxy) has become an important management objective in itself since his original paper. Given this background, $F_{35\%}$ does not, *a priori*, seem well suited to management of an invertebrate fishery that focuses on very large, mature, newly-moulted males with females excluded and both small and "dirty" males (those mature individuals with shell fouling developed after their terminal moult) not preferred by the fishery. A management regime designed to optimize a fishery with that sort of selectivity seems to me to be very unlikely to be similar to one based on $F_{35\%}$ "backed off a bit" (i.e., taking account of the buffer between the OFL and the ABC). As an example, in a rock lobster fishery in New Zealand where much of the mature biomass (male and female) is below the minimum legal size, exploitation targets based on SSB can satisfy the sustainability requirements of New Zealand's Harvest Strategy Standard (HSS, see Ministry of Fisheries 2008) but can result in populations with an unacceptably low incidence of harvestable lobsters (Rudd et al. 2021). A comparable situation seems likely for Tanner crabs; exploitation at $F_{35\%}$ would leave few males in the population of the preferred size and condition preferred by the fishers. Potentially, also, the biomass of effectively reproducing males in the population could become low and threaten recruitment success.

Thoughts on alternative approaches

I'm not in a position to recommend a different reference point at this stage, but I can draw attention to the assessment and management approaches that have been used in New Zealand for the red rock lobster, *Jasus edwardsii*, fishery which has many parallels with the EBS crab fisheries. Rock lobsters are taken commercially using pots, although non-commercial fishers can use other methods, primarily diving. Both males and females are harvested but ovigerous females may not be taken, nor may any below minimum legal sizes or soft (immediate post-moult) individuals. As with Tanner crabs, rock lobsters in some areas become mature before they grow to the minimum legal size and recruit to the fishery, so there is a mismatch between maturity and selection schedules and a significant proportion of the spawning biomass may be protected from all harvesting.

Purely as an example, I will briefly describe the reference points and harvest control rules use for one of the rock lobster stocks, CRA 2. Length-based, integrated stock assessment models have been used for many years (e.g., Starr et al. 2014, Webber et al. 2018, Rudd et al. 2023), fitting to standardized CPUE, tag recaptures, and fishery compositions; there are no fishery-independent biomass indices or compositions. Using the 2013 stock assessment model as an operating model, Starr et al. (2014) developed a range of candidate harvest control rules based on fleet-wide CPUE and tested their performance using a management strategy evaluation and a wide range of performance criteria related to sustainability and fishery outcomes. Between 2013 and 2017, the control rule chosen by the Minister of Fisheries was in force and used to generate a Total Allowable Commercial Catch (TACC) for the Minister to consider for the coming year. The operation of the rule is shown in Figure 6.

Following the adoption of new electronic reporting mechanisms in New Zealand, it became impossible to calculate the CPUE indices needed to drive the control rule. In 2018, therefore, Webber et al. (2018) developed new, interim model-based reference biomass levels for the CRA 2 stock using a method described more fully by Rudd et al. (2021). This assessment was updated in 2022 using the same model-based approach to target reference points based on estimated autumn-winter vulnerable biomass (which includes both sexes but excludes soft-shell lobsters and berried females, and is adjusted to account for gear selectivity and vulnerability). The status of the stock is judged against these model-based reference points and also against limit reference points based on (depletion in) spawning stock biomass of females in the autumn-winter season before fishing (Fisheries New Zealand 2024, Figure 7). This has both parallels and contrasts with the Tanner crab fishery and reference points and these would need to be thought through when designing alternatives for Tanner crabs.

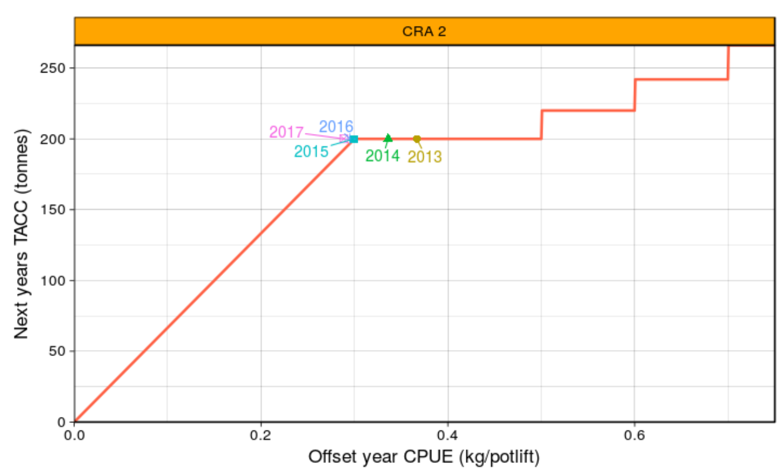

Figure 6: Example CPUE-based harvest control rule for New Zealand rock lobster (*Jasus edwardsii*, CRA 2 stock) used 2013–2017. The rule was developed using MSE. This approach has been superseded by model-based estimates of reference (target) biomass.

A brief description of the method of estimating the target reference point was given by Fisheries New Zealand (2024): "*An interim target reference level for the population vulnerable to fishing (B_R) was developed using the stock assessment model to project the vulnerable biomass through simulation. Starting with the CRA 2 base case model, projections were made across a range of fixed catch and fixed exploitation rate (U) levels. These simulations were for 30 years across the 1000 MCMC posterior samples, with each simulation replicate drawing random recruitment deviates from a lognormal distribution based on the mean, standard deviation, and autocorrelation of model-estimated recruitment deviates over the period 1993–2018. These are the years that span biological sampling in this model which … is the period used for resampling deviates and for basing the estimate of SSB_{0_data}. Only the size-limited catches changed in these projections, with the non-size-limited catch (sum of illegal and customary catch) held constant at the 2021 level. Rules that went below 20% SSB_{0_data} more than 5% of the time for the final 20 simulation years were rejected. The constant catch rules were further constrained by the requirement that at least 99% of the expected catch had to be taken from at least 95% of years in the final 20 simulation years. The vulnerable biomass associated with the fixed catch and fixed U levels that maximised the catch for each harvest strategy while satisfying the constraints over the final 20 simulation years were then averaged to obtain the interim target reference level for CRA 2.*". It can be seen that the target reference point has to satisfy sustainability and operational criteria as well as maximizing catch and, although I am not completely familiar with the US system of setting OFL and ABC (especially legal requirements that might preclude some approaches), I believe a similar simulation method could be used to develop reference points for Tanner crab.

Perhaps as something of a footnote here, the empirical / simulation approach taken to reference points for rock lobsters is somewhat novel so Fisheries New Zealand commissioned an independent review in 2024 (de Lestang et al. 2024). That review panel endorsed the stock assessment modelling approach upon which the simulations are predicated but had a number of reservations about the empirical approach to target reference points and the use of vulnerable biomass rather than spawning biomass. They considered that a better justification of the reasons for, and advantages of, its adoption over the better-known default proxy (in the New Zealand system, see Ministry of Fisheries 2008 and 2011) of 40% B_0. However, they also made the point that accounting for the objectives that target reference points are intended to meet would make using a simulation-based vulnerable biomass target reference point less arbitrary than selecting 40% B_0 simply because it was the default. Again, there are many parallels with the discussion about reference points for Tanner crabs that we had at the review meeting, and I think this approach could be explored for Tanner crabs (and other *Chionoecetes* crabs).

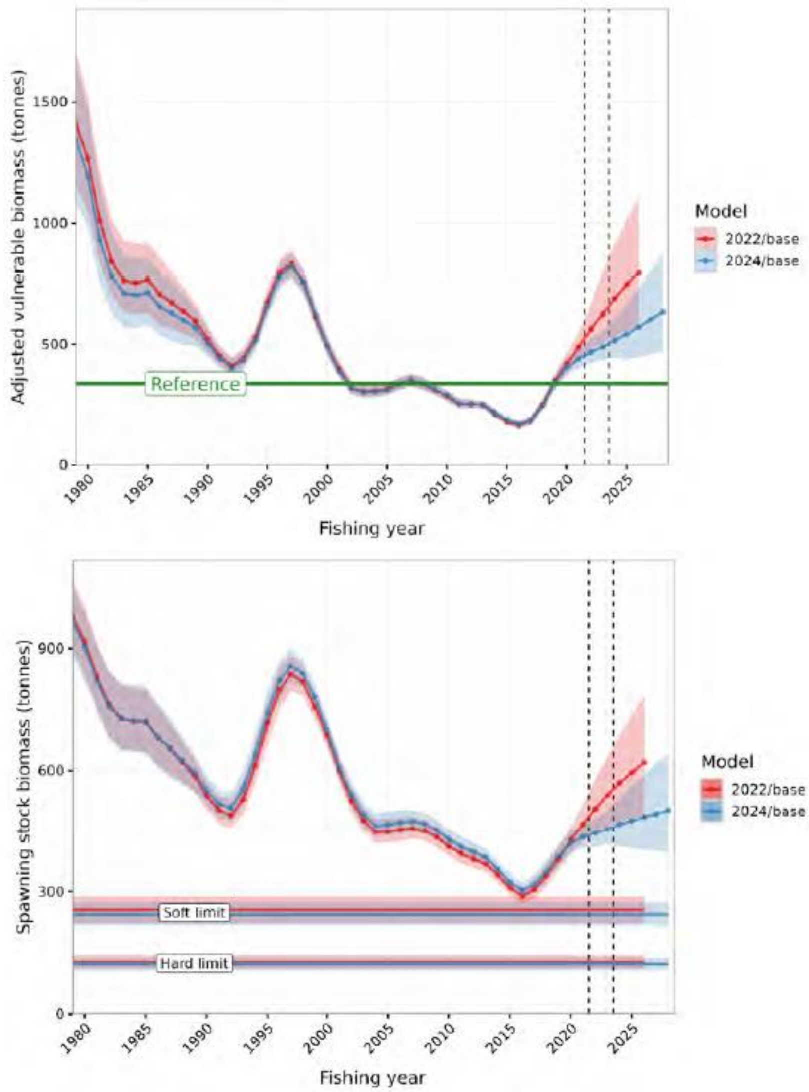

Figure 7: (from MPI 2024, p344). Top: posterior distribution of the Autumn-Winter adjusted vulnerable biomass trajectory for the CRA 2 stock and projections from the 2022 full assessment and 2024 rapid update assessment. The B_R interim biomass reference target is shown as a solid green line. Bottom: posterior distribution of the female spawning stock biomass trajectory compared with the soft (20% SSB_0) and hard (10%SSB_0) biomass limit reference points. In both plots, variable shading indicates the 50% and 90% credible intervals and the solid line indicates the median. The vertical dashed lines indicate the transition to the projected years.

In summary, I don't believe that a reference point of $F_{35\%SPR}$ (or $F_{40\%SPR}$) is a useful or defensible approach for this combination of biology and fishery where only a small fraction of the population has value to the fishers. I can't recommend another specific reference point, but I believe a more empirical, simulation-based approach, at least to a target reference point, would be more fruitful and, potentially, provide management advice that is more directly relevant. A significant amount of work would be needed to develop such an approach, including finalizing an operating model (presumably based on the stock assessment model), identifying the important ecological, sustainability and operational performance criteria, and running and summarizing the simulations. As for limit reference points, I believe these should remain strictly biologically-based. Whether they reference male or female mature biomass, I am less certain. The rationale for using mature male biomass is grounded in the fact that this is the portion of the population that is subject to depletion by the fishery. However, given the large and sometimes rapid fluctuations in the Tanner crab stock, mature female biomass could also fall to low levels independent of fishing and even when male mature biomass is high (Figure 8, especially 2013 and 2014). That being the case, it would be prudent to consider female as well as male biomass in some way. I believe a vital part of the background work needed to underpin development of new limit reference points for Tanner crabs should be an assessment of what led to population crashes for some “stocks” in Alaska while other “stocks” persisted. I heard about this for the first time at the review meeting and I’m not sure what information is available to review or re-analyse.

ToR4: Need for two-area models

This ToR called for an assessment of whether a two-area model for Tanner crab should be recommended, and what, if any, additional work could inform this decision.

In general, spatially structured stock assessment models may be preferred if there are substantial differences in the productivity and/or exploitation of different parts of a wider population and these features could lead to over- and under-exploitation within the whole area (e.g., Cardinale et al. 2023, Goethel et al. 2024). These criteria are more often and more obviously met for invertebrates than for mobile finfish (Caddy 1975 and much work since). However, there are also downsides related to data needs, bias, model complexity, dimensionality and fitting that should not be underestimated. Goethel et al. (2024) concluded from their blind trials that spatial models estimated higher population scale than single-region models within a given assessment platform, although the impact on estimated depletion was smaller and less consistent. They also found that balancing parsimony and complexity trade-offs was difficult.

The trawl survey results presented to the review panel showed quite clearly that there is a substantial amount of spatial pattern in the Tanner crab population and that this has varied between years. There is some suggestion that there has been a northward drift of the population’s “center of gravity” and that the most recent year is very unusual, both being potential responses to climate change. However, my view, based on the material provided and the discussions at the review meeting, is that the 2025 results are not particularly noteworthy in the context of what appears to be a naturally variable stock or population subject to spatially varying drivers of productivity (recruitment success, growth, and/or mortality). My strongest “evidence” for this view, which I admit is relatively weak, is the broad synchrony of peaks in biomass in the eastern and western areas and the similarity of survey results for 1982 and 2024/25.

In summary, I am not expert in this field but my background in the population dynamics and assessment of invertebrate populations, plus what I have read, lead me to believe that developing a multi-area model for EBS Tanner crab is not a priority at this stage. The comprehensive broad-scale data generated by the NMFS trawl survey provide plenty of spatial data that could be used to inform management at finer spatial scales if this is thought desirable to avoid local depletion or unbalanced sex ratios. Given the grid-based design, the same survey data could be readily disaggregated for a multi-area model if one is required in the future to underpin more precise advice on depletion in different areas and any necessary catch reductions or spreading. This seems a low priority relative to work on the current reference points which lead to catch advice that seems so high as not to be useful for the managers.

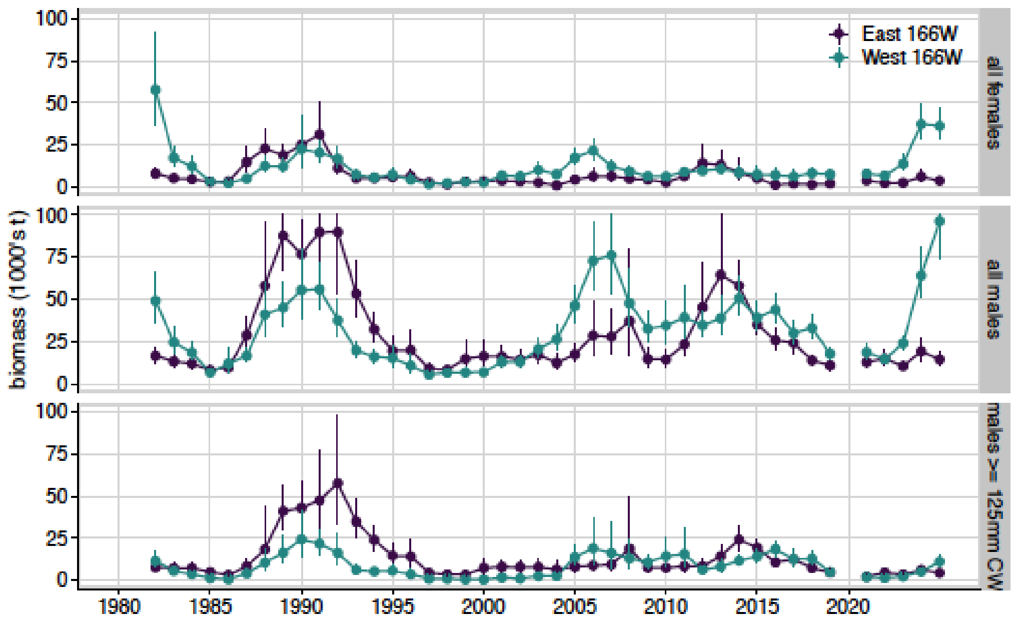

Figure 8: Plots of female, male, and large male biomass from trawl survey stations east and west of 166 °W in the EBS. Note broad synchrony of peaks, spatially, and the similarity of results at the start and end of this time series.

ToR5: Treatment of hybrid *Chionoecetes* crabs in the assessment

This ToR called for an assessment of how hybrid *Chionoecetes* crab should be treated in the assessment and in reference point and harvest control rule calculations, and advice on additional work that could better inform these decisions.

Based on the material provided, the estimated proportion and abundance of *Chionoecetes* hybrids in the latest NMFS trawl surveys is at a historical high. However, there seems to be quite a bit of subjectivity in the morphological identification of hybrids and Urban et al. (2002) reported high variance in the accuracy of experts and trainees. We heard in the review meeting that, probably as a consequence of the subjectivity, it is possible that there may be "drift" in detectability over time as on-board staff become more aware of hybrids as an issue. It is also conceivable that many more hybrids exist that are morphologically very similar to, or indistinguishable from, "pure" snow or Tanner crabs, particularly where the two species overlap (Johnson 2019, Slater et al. 2024, and references in both). My own experience with such "cryptic" hybrids is in freshwater cyprinid fishes (and is very dated), but natural hybridization occurs in many shrimp, lobster, crayfish and crab species (e.g., Syafaat et al. 2024) and hybrids can often be identified with certainty only using genetic analysis. Given the recent advances in rapid genetic assessments (e.g., Edwards et al. 2025), a broad-scale molecular study of the prevalence and distribution of hybrids might be worthwhile but, for me, it would be only a low-moderate priority unless the proportion estimated from the survey continues to rise.

A sustainability issue could arise if a high proportion of hybrid crabs are included in the stock assessment as part of the spawning biomass and they are not fully functional spawners; clearly this could lead to overestimation of the effective spawning biomass. Conversely, if hybrids are excluded, then the biomass available for fishing will be underestimated. It is known that *Chionoecetes* hybrids can spawn and produce viable and reproductively active offspring (Johnson et al. 2019), although we were told that some data suggest that they may not produce as many offspring as “pure” crabs so they may not be “fully functional”. The treatment of hybrids in the stock assessment and management decision-making would depend, in part, on the appetite for these opposing risks. Even at the current (historically high) proportion of hybrids, however, including all hybrids (regardless of legal allocation to species) in the Tanner crab assessment model did not greatly affect the conclusions (Figure 9). That being the case, unless the proportion of hybrids continues to increase markedly, the treatment of hybrids in the stock assessment is of limited relevance, and a low priority compared with other work. Plenty of information could be derived from the NMFS trawl surveys to provide for additional catch of hybrids through harvest strategies, control rules, or *ad hoc* measures if required.

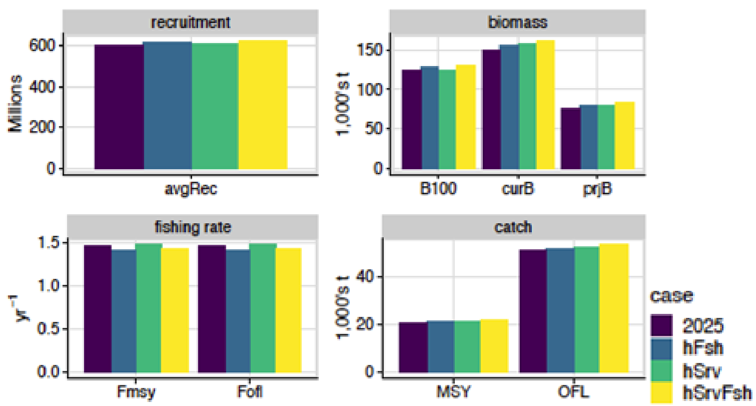

Figure 9: Comparison of values for various management reference quantities from the 2025 assessment model (“2025”) and the models including hybrid survey (“hSrv”), hybrid fishery (“hFsh), and hybrid survey and fishery (”hSrvFsh”) data. (Copy of Figure 37 from Stockhausen 2026b).

In summary, I don’t believe the recent high estimates of *Chionoecetes* hybrids in the NMFS survey are a cause for much concern, although monitoring should continue. No sustainability risk is generated by the current practice of excluding hybrids from the Tanner crab assessment and I believe (contingent on the intricacies of the US system with which I am not very familiar) additional catch targeted on hybrids could be provided for in other ways.

Other thoughts on the stock assessment model

The ToR for this review did not call for comments on stock assessment model itself, but I believe there are aspects of the model that are relevant to the specified ToR, especially ToR2 (catchability and selectivity) and, especially, ToR3 (reference points and harvest control rules).

A striking feature of the model fits for all Tanner crab stock assessments I’ve looked at (those for 2000, 2003, and 2025) is the systematic inability of the model to fit the magnitude of changes in biomass estimates from the NMFS trawl survey. It is not uncommon (and, in fact, desirable) for models to smooth through unrealistically large changes in biomass from surveys that might stem from changes to catchability or availability among surveys, but the pattern for Tanner crabs is qualitatively different; the biomass trajectory fails to duplicate the magnitude of any the convincing-looking survey peaks resulting in 3–6 consecutive and very large residuals (Figure 10) for each peak.

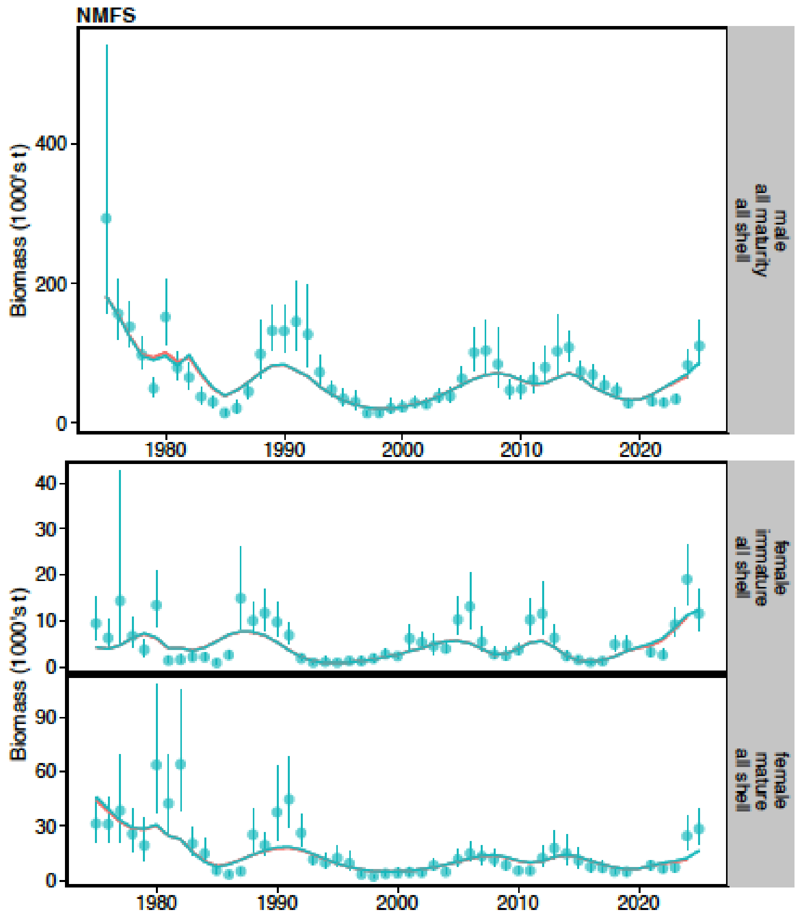

Figure 10: Fit of the stock assessment model to the NMFS trawl survey time series. For all categories, the model does not replicate the magnitude of observed peaks and troughs in biomass spanning several years.

My immediate reaction to this was to suspect that the survey biomass indices had been under-weighted in the model compared with the compositions, or that priors and penalties associated with M are too influential. This view was reinforced for me by the thought that model cohorts recruit at a size of 25 mm carapace width and high growth variability across the several years between settlement and this size could “smear” recruiting year-classes across the compositions, confusing the recruitment pattern. After reading about the stock assessment in more detail, I didn’t see any obvious issues with the data set weightings (although they seemed slightly *ad hoc*) but undue influence of the compositions and priors remain a possibility for me. Were this a review of the stock assessment model itself, I would definitely want to see the impact of radically different treatment of both and some likelihood profiling to work out what was driving the patterns. I imagine, however, that a much higher estimate of M in the model, allied with higher σR, would be required to allow the steep increases and declines, and these may not be reasonable given the biology of the species.

Other potential drivers that occurred to me were temporally autocorrelated changes in survey catchability (unlikely) and density dependence in mortality and/or growth. I was not the first to think of the latter and it came up at the review meeting. Low M at low biomass would facilitate relatively rapid increases in biomass and high M setting in during the peaks would drive steeper declines than the current model can accommodate. Further evidence for substantial density-dependent factors at work in this population comes from the empirical stock-recruit relationships provided during the review meeting (Figure 11). These clearly imply domed stock-recruit relationships which, in turn, imply density-dependent feedback.

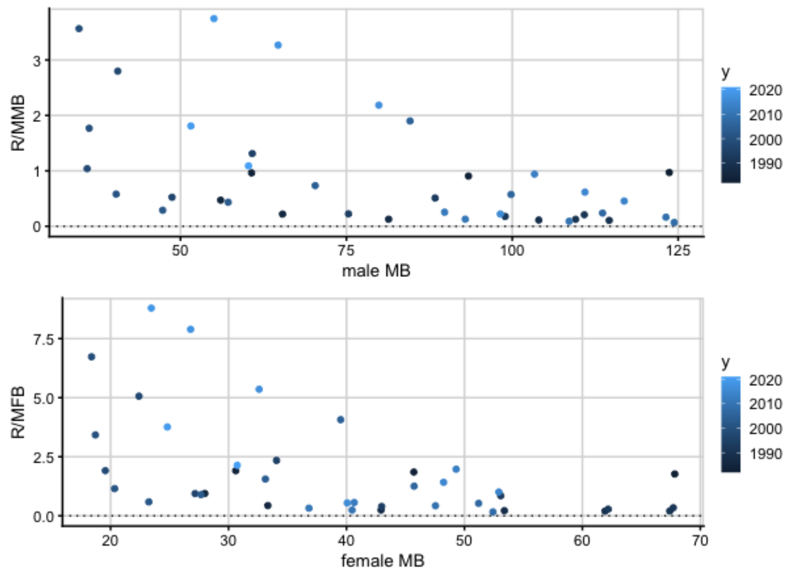

Figure 11: Relationship between mature biomass (top = males, bottom = females) and estimated recruitment of new individuals to the stock assessment model 4 years later. A domed rather than asymptotic relationship is apparent in both.

I think this would be worth exploring further. A first step might be to eliminate weighting issues and survey catchability (relevant to ToR2) as possibilities; both could be done without extensive model development. Building models (in GMACS) with density-dependence in M and/or recruitment would take longer, but would have implications for reference points and harvest control rules under ToR3, especially if a simulation approach to estimating target reference points is taken. Another, somewhat heuristic, approach might be to allow M to vary annually in a model and explore relationships of the M deviates with external environmental drivers and biomass (of males, females, or all likely predatory size classes?). This might lead on to a model involving one or more environmental drivers and/or density-dependence. A state-space approach was discussed at the review meeting but this is not really my area of expertise and I won't develop the idea here.

Conclusions and Recommendations

I believe the stock assessment team, led by Buck Stockhausen, have done a very sound and very thorough job with the Tanner crab assessment and it was a pleasure working with them on this review. I have lots of views on how things could be done differently, but I offer these as ideas to try rather than, necessarily, superior approaches. I summarise those thoughts by ToR below.

ToR1 (GMACS framework): I believe the pros of moving to the GMACS framework now greatly outweigh the cons and I see no concerns that would preclude using GMACS.

ToR2 (SBS analysis): I believe the methods used to estimate the catchability and selectivity of the NMFS survey gear based on the side-by-side experiments are broadly sound and I believe that, noting some ongoing uncertainty, the information has been included in the stock assessment model appropriately. Although I have some reservations about the area swept assumptions in the SBS experiments, there is a strong and consistent signal on the different performance of NMFS vs Nephrops gear in the SBS data which means that quite large changes in population scale in the assessment model are to be expected and defensible. The model is, therefore, suitable for providing management advice. I believe additional work could be done to bolster the estimates of catchability and selectivity of the survey gear, potentially making those estimates more believable for stakeholders. Such work could include some or all of exploring "lift-off lags" in the two surveys, pooling data for all *Chionoecetes* crabs, doing depletion-based experiments, and fitting models direct to the data instead of to ratios, and using logistic, Richards or normal curves rather than GAMs or spline-based models. This work may increase or decrease the NMFS survey catchability estimates and lead to smaller or larger changes in population scale than exhibited by the current stock assessment model.

ToR3 (reference points): I don't believe that a reference point of $F_{35\%SPR}$ (or $F_{40\%SPR}$) is a useful or defensible approach for this combination of biology and fishery where only a small fraction of the population has value to the fishers. I believe a more empirical, simulation-based approach, at least to a target reference point, would be more fruitful and, potentially, provide management advice that is more directly relevant. Although this would entail a significant amount of work including finalizing an operating model, identifying the important ecological, sustainability and operational performance criteria, and running and summarizing the simulations, I think this is a high priority. I believe limit reference points should remain strictly biologically-based but I am less clear on whether they should reference male or female mature biomass. The rationale for using mature male biomass is grounded in the fact that this is the portion of the population that is subject to depletion by the fishery. However, given the large and sometimes rapid fluctuations in the Tanner crab stock, mature female biomass could also fall to low levels independent of fishing and even when male mature biomass is high. That being the case, it would be prudent to consider female as well as male biomass in some way. I believe a vital part of the background work needed to underpin development of new limit reference points for Tanner crabs should be an assessment of what led to population crashes for some "stocks" in Alaska while others persisted.

ToR4 (multi-area modelling): I don't believe that developing a multi-area model for EBS Tanner crab is a priority at this stage. The comprehensive broad-scale data generated by the NMFS trawl survey provide plenty of spatial data that could be used to inform management at finer spatial scales if this is thought desirable to avoid local depletion or unbalanced sex ratios. Given the grid-based design, the same survey data could be readily disaggregated for a multi-area model if one is required in the future to underpin more precise advice on depletion in different areas and any necessary catch reductions or spreading. I think this is a low priority relative to work on reference points.

ToR5 (hybrids): I don't believe the recent high estimates of *Chionoecetes* hybrids in the NMFS survey are a cause for much concern, although monitoring should continue. No sustainability risk is generated by the current practice of excluding hybrids from the Tanner crab assessment and I believe additional catch targeted on hybrids could be provided for in other ways.

General: The current stock assessment model has a striking (and intriguing) inability to reproduce very believable peaks in the survey biomass index. The reasons for this should be investigated because they may have significant implications for how the survey data are included in the model and how reference points might be calculated, especially if a simulation-based approach is taken to estimating target reference points. I suggest exploring data weighting, survey catchability issues, potential environmental drivers, and density dependence.

References

Bian, R.; Williams, J.R.; Smith, M.; Tuck, I.D. (2012). Modelling dredge efficiency for the Coromandel and Northland scallop fisheries. Final Research Report for project SAP200913. Fisheries New Zealand, Wellington, New Zealand. 46 p.

Caddy J.F. (1975). Spatial model for an exploited shellfish population, and its application to the Georges Bank scallop fishery. *Journal of the Fisheries Research Board of Canada* **32(8)**: 1305–1328. doi: 10.1139/f75-152

Clark, W.G. (1991). Groundfish exploitation rates based on life history parameters. *Canadian Journal of Fisheries and Aquatic Sciences* **48(5)**: 734–750.

Clark, W.G. (1992). of Spawning Biomass Per Recruit. pp 233–246 in: Kruse, G.; Eggers, D; Marasco, R.; Pautzke, C.; Quinn, T. (eds). *Proceedings of the International Symposium on Management Strategies for Exploited Fish Populations*, Lowell-Wakefield Fisheries Symposium, October 21–24, 1992, Anchorage, Alaska.

Clark, W.G. (2002). F35% revisited ten years later. *North American Journal of Fisheries Management* **22(1):** 251–257.

Cryer, M.; Parkinson, D.M. (2006). Biomass surveys and stock assessments for the Coromandel and Northland scallop fisheries, 2005. *New Zealand Fisheries Assessment Report 2006/34*. 54 p.

Delargy, A.J.; Blackadder, L.; Bloor, I.; McMinn, C.; Rudders, D.B.; Szostek, C.L.; Dobby, H.; Kangas, M.; Stewart, B.D.; Williams, J.R.; Stokesbury, K.D. (2023). A global review of catch efficiencies of towed fishing gears targeting scallops. *Reviews in Fisheries Science & Aquaculture* **31(3)**: 296–319.

Edwards, M.; Ellis, C.; Batista, F. (2025). Rapid and cost-effective differentiation of the lobsters *Homarus americanus*, *H. gammarus* and their F1 hybrids using DNA-based methods. *Arthropoda* **3(4)**: p.15.

Engås, A.; Godø, O.R.; Jørgensen, T. (2000). A comparison between vessel and trawl tracks as observed by the ITI trawl instrumentation. *Fisheries Research* **45(3)**: 297–301.

Fisheries New Zealand (2024). *Fisheries Assessment Plenary, November 2024*: stock assessments and stock status. Compiled by the Fisheries Science Team, Fisheries New Zealand, Wellington, New Zealand. 691 p.

Fulanda, B.; Ohtomi, J. (2011). Determination of effective tow duration for estimation of fish and shellfish abundance in deeper-water bottom trawl surveys. *Fisheries Science* **77**: 487–495.

Goethel, D.R.; Berger, A.M.; Hoyle, S.D.; Lynch, P.D.; Barcelo, C.; Deroba, J.; Ducharme-Barth, N.D.; Dunn, A.; Fu, D.; Izquierdo, F.; Marsh, C.; Xu, H.; Correa, G.M; Langseth, B.J.; Maunder, M.N.; McKenzie, J.R.; Methot, R.D.; Vincent, M.T.; A'mar, T.; Cardinale, M.; Cousido-Rocha, M.; Davies, N.M.; Hampton, J.; Minte-Vera, C.; Urtizberea, A. (2024). 'Drivin' with your eyes closed': Results from an international, blinded simulation experiment to evaluate spatial stock assessments. *Fish and Fisheries* **25(3)**: 471–490.

Hennen, D.R.; Jacobson, L.D.; Tang, J. (2012). Accuracy of the patch model used to estimate density and capture efficiency in depletion experiments for sessile invertebrates and fish. *ICES Journal of Marine Science* **69(2)**: 240–249.

Johnson, G.M. (2019). Genetic diversity and population genetic structure of Tanner crab *Chionoecetes* bairdi in Alaskan waters. Thesis submitted in partial fulfilment of the requirements for the degree of Master of Science in Fisheries. University of Alaska, Fairbanks. 71 p.

de Lestang, S.; Haddon, M.; Hoyle, S. (2024). Review of red rock lobster stock assessment modelling and the determination of management reference points. *New Zealand Fisheries Science Review 2024/01*. 28 p.

Ministry of Fisheries (2008). Harvest Strategy Standard for New Zealand's fisheries. https://www.mpi.govt.nz/dmsdocument/728-Harvest-Strategy-Standard-for-New-Zealand-Fisheries.

Ministry of Fisheries (2011) Operational Guidelines for New Zealand's Harvest Strategy Standard. Revision 1. 78 p. https://www.mpi.govt.nz/dmsdocument/19706-OPERATIONAL-GUIDELINES-FOR-NEW-ZEALANDS-HARVEST-STRATEGY-STANDARD

Munroe, D.; Borsetti, S.; Sheehan, A.; Piper, S.; Morson, J.; Hennen, D. (2025). Estimating efficiency of a scientific clam survey dredge: strategies for an improved experimental approach. *North American Journal of Fisheries Management* **45(5):** 979–986.

Poussard, L.M.; Powell, E.N.; Hennen, D.R. (2021). Efficiency estimates from depletion experiments for sedentary invertebrates: evaluation of sources of uncertainty in experimental design. *Fisheries research* **234**: p.105806.

Rago, P.J.; Weinberg, J.R.; Weidman, C. (2006). A spatial model to estimate gear efficiency and animal density from depletion experiments. *Canadian Journal of Fisheries and Aquatic Sciences* **63(10):** 2377–2388.

Rudd, M.B.; Pons, M.; Webber, D.N.; Starr, P.J.; Roberts, J.; Goeden Z.D. (2023). The 2022 stock assessment of red rock lobsters (*Jasus edwardsii*) in CRA 2. *New Zealand Fisheries Assessment Report 2023/43*. 108 p.

Rudd, M.B; Webber, D.N.; Starr, P.J. (2021) Model-based reference levels for New Zealand red rock lobster (*Jasus edwardsii*). *New Zealand Fisheries Assessment Report 2021/81*. 42 p. https://fs.fish.govt.nz/Doc/25019/FAR-2021-81-Rock-Lobster-Model-Based-Reference-Levels-4218.pdf

Saud, M.A.J. (2007). Sensitivity of the F35% target fishing rate to variation in the maturity schedule. *Journal of Fisheries International* ***2*****(4)**: 284–291.

Slater, L.M.; Gaeuman, W.; Cheng, W.; Kruse, G.H.; Habicht, C.; Pengilly, D. (2024). Molecular evaluation of the mating dynamics of snow crab (*Chionoecetes opilio*) in the Eastern Bering Sea. *Ecology and Evolution* **14(10):** p.e70416.

Smith, S.; Hart, D.; Haddon, M. (2016). Review of New Zealand's scallop fishery stock assessment data and methods. *New Zealand Fisheries Science Review 2016/1*. 25 p.

Somerton, D., Weinberg, K., and Goodman, S. (2010?). Snow crab selectivity by the NMFS trawl survey. Unpublished NOAA report. 23 p.

Somerton, D.A.; Weinberg, K.L.; Goodman, S.E. (2013). Catchability of snow crab (*Chionoecetes opilio*) by the eastern Bering Sea bottom trawl survey estimated using a catch comparison experiment. *Canadian Journal of Fisheries and Aquatic Sciences*, **70(12)**: 1699–1708.

Starr, P.J.; Haist, V.; Breen, P.A.; Edwards, C.T.T.E. (2014). The 2013 stock assessment of red rock lobsters (*Jasus edwardsii*) in CRA 2 and development of management procedures. *New Zealand Fisheries Assessment Report 2014/19*. 76 p.

Stockhausen, W.T. (2020). 2020 Stock assessment and fishery evaluation report for the Tanner crab fisheries of the Bering Sea and Aleutian Islands regions. Alaska Fisheries Science Center, Seattle. 183 p.

Stockhausen, W.T. (2023). 2023 Stock assessment and fishery evaluation report for the Tanner crab fisheries of the Bering Sea and Aleutian Islands regions. Alaska Fisheries Science Center, Seattle. 339 p.

Stockhausen, W.T. (2025). 2023 Stock assessment and fishery evaluation report for the Tanner crab fisheries of the Bering Sea and Aleutian Islands regions. Alaska Fisheries Science Center, Seattle. 467 p.

Stockhausen, W.T. (2026a). Side-by-side survey catchability studies: cooperative research to estimate survey catchability for Tanner crab in the Eastern Bering Sea. DRAFT NOAA Technical Memorandum. 118 p.

Stockhausen, W.T. (2026b). 2026 Tanner crab assessment considerations. Document dated 2026-05-04. 151 p.

Syafaat, M.N.; Abualreesh, M.H.; Yatim, N.I.; Fazhan, H.; Waiho, K.; Ma, H.; Okomoda, V.T.; Ikhwanuddin, M. (2024). Interspecific hybridization of decapod crustacean species with commercial interest—A review. *Reviews in Aquaculture* **16(2)**: 741–758.

Tremblay, M.J.; Smith, S.J.; Robichaud, D.A.; Lawton, P. (2006). The catchability of large American lobsters (*Homarus americanus*) from diving and trapping studies off Grand Manan Island, Canadian Maritimes. *Canadian Journal of Fisheries and Aquatic Science* **63**: 1925–1933.

Tuck, I.D.; Brown, S. (2008). Survey of scallops and oysters in Golden Bay, Tasman Bay, and the Marlborough Sounds, 2008. NIWA Client Report: NEL2008-022. July, 2008.

Urban, D.; Pengilly, D.; Jadamec, L.; Byersdorfer, S. C. (2002). Testing carapace morphology characteristics for field identification of *Chionoecetes* hybrids. In A. J. Paul, E. G. Dawe, R. Elner, G. S. Jamieson, G. H. Kruse, R. S. Otto, & B. Sainte-Marie (Eds.), Crabs in coldwater regions: biology, management and economics. *Lowell Wakefield Fisheries Symposium Series*. Alaska Sea Grant College Program University of Alaska, Fairbanks (pp. 97–113).

Vølstad, J.H.; Sharov, A.F.; Davis, G.; Davis, B. (2000). A method for estimating dredge catching efficiency for blue crabs, *Callinectes sapidus*, in Chesapeake Bay. *Fishery Bulletin* **98**: 410–420.

Wallace, J.R.; West, C.W. (2006). Measurements of distance fished during the trawl retrieval period. *Fisheries Research* **77(3)**: 285–292.

Webber, D.N.; Starr, P.J.; Haist, V.; Rudd, M.; Edwards, C.T.T. (2018a). The 2017 stock assessment and management procedure evaluation for rock lobsters (*Jasus edwardsii*) in CRA 2. *New Zealand Fisheries Assessment Report 2018/17*. 87 p.

West, C.W.; Wallace, J.R. (2000). Measurements of distance fished during the trawl retrieval period. ICES CM 2000/K:37. 14 p.

Wilberg, M.J.; Robinson, J.M.; Rains, S.A.; Humphrey, J.L.; Lipcius, R.N. (2013). Effects of location errors on estimates of dredge catchability from depletion-based methods. *Fisheries Research* **148**: 1–8.

Yochum, N.; Dupaul, W.D. (2008). Size-selectivity of the northwest Atlantic sea scallop (*Placopecten magellanicus*) dredge. *Journal of Shellfish Research* **27(2)**: 265–271.

Appendix 1: Bibliography of materials provided for review

The following materials were provided before or during the review meeting although I read quite a few other reports, papers and documents to inform my review (see list of cited references above).

Reports and papers

Anon (2025). BSAI Crab SAFE Introduction dated OCTOBER 2025. 10 p.

Heller-Shipley, M.A., Stockhausen, W.T., Daly, B.J., Punt, A.E. and Goodman, S.E., (2021). Should harvest control rules for male-only fisheries include reproductive buffers? A Bering Sea Tanner crab (*Chionoecetes bairdi*) case study. *Fisheries Research, 243*, p.106049.

Stockhausen, W.T. (2026). 2026 Tanner Crab Assessment Considerations, dated 4 May 2026. 151 p.

Stockhausen, W.T. (2026). Addendum to Transition from TCSAM02 to GMACS: 1 function call, dated 11 May 2026. 7 p.

Stockhausen, W.T. (2026). Side-by-Side Survey Catchability Studies: Cooperative Research to Estimate Survey Catchability for Tanner Crab in the Eastern Bering Sea. DRAFT NOAA Technical Memorandum. 118 p.

Stockhausen, W.T. (2026). Tanner crab Stock Structure and Spatial Considerations, dated 25 May 2026. 13 p.

Zacher, L. S., Hennessey, S. M., Richar, J. I., Fedewa, E. J., Ryznar, E. R., and Litzow, M. A. (2026). The 2025 eastern and northern Bering Sea continental shelf trawl surveys: Results for commercial crab species. U.S. Department of Commerce, NOAA Technical Memorandum NMFS-AFSC-513, 299 p.

Zaleski, M., T. S. Smeltz, S. Gardiner, J. L. Pirtle, and G. A. Harrington. (2024). 2022 Evaluation of Fishing Effects on Essential Fish Habitat. NOAA Technical Memorandum NMFS-F/AKR-29, 212 p. doi: 10.25923/c2gh-0w03

Presentations

Stockhausen, W.T. (2026). Transition To GMACS. 31 p.

Stockhausen, W.T. (2026). Estimated Selectivity Using Side-by-Side Haul Data. 21 p.

Stockhausen, W.T. (2026). Management and Tier Considerations. 12 p.

Stockhausen, W.T. (2026). Spatial Considerations. 11 p.

Stockhausen, W.T. (2026). Hybrid considerations. 11 p.

Stockhausen, W.T. (2026). Assignments, Day 1. 4 p.

Stockhausen, W.T. (2026). Assignments, Day 2. 3 p.

Other

All R-code was provided on a github repository, together with many of the graphics from the reports. A figure showing snow crab distributions for the years from 1976 to 2023 was also circulated.

Appendix 2: CIE Performance Work Statement

Performance Work Statement (PWS)
National Oceanic and Atmospheric Administration (NOAA)
NOAA Fisheries
Center for Independent Experts (CIE) Program
External Independent Peer Review

Review of Eastern Bering Sea Tanner Crab Assessment
June 9 - 11, 2026

Background

NOAA Fisheries is mandated by the Magnuson-Stevens Fishery Conservation and Management Act, Endangered Species Act, and Marine Mammal Protection Act to conserve, protect, and manage our nation's marine living resources based upon the best scientific information available (BSIA). NOAA Fisheries science products, including scientific advice, are often controversial and may require timely scientific peer reviews that are strictly independent of all outside influences. A formal external process for independent expert reviews of the agency's scientific products and programs ensures their credibility. Therefore, external scientific peer reviews have been and continue to be essential to strengthening scientific quality assurance for fishery conservation and management actions.

Scientific peer review is defined as the organized review process where one or more qualified experts review scientific information to ensure quality and credibility. These expert(s) must conduct their peer review impartially, objectively, and without conflicts of interest. Each reviewer must also be independent from the development of the science, without influence from any position that the agency or constituent groups may have. Furthermore, the Office of Management and Budget (OMB), authorized by the Information Quality Act, requires all federal agencies to conduct peer reviews of highly influential and controversial science before dissemination, and that peer reviewers must be deemed qualified based on the OMB Peer Review Bulletin standards[2].

Scope

The goal of this review is to gain external expert review and input on several issues with respect to the Bering Sea and Aleutian Islands Tanner crab assessment, including:

1. transition of the Tanner crab assessment to the Generalized Model for Assessing Crustacean Stocks (GMACS) framework,
2. methodology for incorporation of data collected in side-by-side selectivity experiments into the Tanner crab assessment,
3. the appropriateness of currently used reference points for Tanner crab (and potential

[2] https://www.whitehouse.gov/wp-content/uploads/legacy_drupal_files/omb/memoranda/2005/m05-03.pdf

alternatives),
4. the feasibility and need for development of a two-area model for Tanner crab
5. how to treat hybrid snow/Tanner crab in the assessment and management process

External expert review for this stock is necessary and timely because Tanner crab abundance is currently increasing strongly and hybridization with snow crab has reached levels not historically observed. If these trends continue, Tanner crab may become a larger part of the portfolio of fished crab stocks in the Bering Sea than it has been for the past several decades. External review will help ensure the science used in the management of Tanner crab continues to be robust and transparent in the pursuit of management advice that maximizes long-term sustainable catches.

The specified format and contents of the individual peer review report are found in **Annex 1**. The Terms of Reference (ToRs) of the peer and review are included in **Annex 2**. The tentative agenda of this virtual panel review and process review meetings is included in **Annex 3**.

Requirements
NOAA Fisheries requires three reviewers to conduct an impartial and independent peer review in accordance with this Performance Work Statement (PWS), OMB Guidelines, and the ToRs below. Modifications to this Performance Work Statement (PWS) and ToRs cannot be made during the peer review, and the Contracting Officer's Representative (COR) and the CIE contractor shall approve any modifications prior to the peer review. All ToRs must be addressed in each reviewer's report.

Effective review of the issues of concern requires reviewers with several specific skillsets. In-depth experience with the implementation and behavior of *integrated size-structured assessment models* that have complex internal dynamics and are fit to numerous data sources is needed to appropriately review the Tanner crab model. Experience in the *management of crustaceans* is needed to support the analysis and discussion of reference points and harvest control rules. Finally, knowledge of *crustacean biology* is needed given the distinct structure of Tanner crab population dynamics and the resulting implications for management.

The chair, who is in addition to the panel of three reviewers, will not be provided by the CIE. Although the chair will be participating in this review, the chair's participation (e.g., labor and travel) is not covered by this contract.

Tasks for Reviewers
Each CIE reviewer shall complete the following tasks in accordance with the PWS and Schedule of Milestones and Deliverables herein.

<u>Pre-review Background Documents</u>: Two weeks before the peer review, the NOAA Fisheries Project Contact will send (by electronic mail or make available at an FTP site) to the CIE reviewers the necessary background information and reports for the peer review. In the case where the documents need to be mailed, the NOAA Fisheries Project Contact will consult with

the CIE on where to send documents. The CIE reviewers are responsible only for the pre-review documents that are delivered to the reviewer in accordance with the PWS scheduled deadlines specified herein. The CIE reviewers shall read all documents in preparation for the peer review.

Documents to be provided to the CIE reviewers prior to this review include:

Documents for review (links to come):

- Document describing issues for which input is needed
- 2025 Tanner Crab Assessment
- May 2026 Preliminary Tanner crab models and other analyses

Background information:

- Survey tech memo
- previous MSE work (BSFRF, UW, ADFG, AFSC)
- State TAC docs
- Chionoecetes mortality papers
- Hybrid crab background

Panel Review Meeting: The CIE reviewers shall conduct the independent peer review in accordance with the PWS and ToRs and shall not serve in any other role unless specified herein. The CIE reviewers shall attend and actively participate in a professional and respectful manner as members of the meeting review panel, and their peer review tasks shall be focused on the ToRs as specified herein. The NOAA Fisheries Project Contact is responsible for any meeting arrangements (e.g., conference room reservations, video conferencing logistics, etc.).

Chair Summary Report: The reviewers shall assist the Chair of the meeting with contributions to the summary report. Reviewers are not required to reach a consensus.

Contract Deliverables - Independent CIE Peer Review Report: The CIE reviewers shall complete an independent peer review report in accordance with this PWS. The CIE reviewer shall complete the independent peer review according to required format and content as described in **Annex 1**. The CIE reviewer shall complete the independent peer review addressing each ToR as described in **Annex 2**. The tentative agenda of the panel review meeting is attached in **Annex 3**.

Foreign National Security Clearance

When reviewers participate during a panel review meeting at a government facility, the NOAA Fisheries Project Contact is responsible for obtaining the Foreign National Security Clearance approval for reviewers who are non-US citizens. For this reason, the reviewers shall provide requested information (e.g., first and last name, contact information, gender, birth date, passport number, country of passport, travel dates, country of citizenship, country of current residence, and home country) to the Project Contact for the purpose of their security clearance, and this information shall be submitted at least two weeks in advance. For additional information, please see the following link: https://www.commerce.gov/osy/programs/foreign-access-management. All CIE reviewers must have a REAL ID-compliant form of identification to access Federally owned or leased facilities The contractor is required to use all appropriate methods to safeguard Personally Identifiable Information (PII).

Place of Performance
The place of performance shall be the Alaska Fisheries Science Center, Seattle, WA.

Period of Performance
The CIE reviewers' duties shall not exceed **14** days to complete all required tasks.

Schedule of Milestones and Deliverables
The contractor shall complete the tasks and deliverables in accordance with the following schedule:

Schedule	Milestones and Deliverables
Within two weeks of award	Contractor secures participation of the CIE reviewers.
At least two weeks prior to the panel review meeting	Pre-review documents are provided to the reviewers.
June 9 - 11, 2026	The reviewers participate and conduct an independent peer review during the panel review meeting.
Approximately 2 weeks later	Contractor receives draft reports
Within 3 weeks of receiving draft reports	Contractor submits final reports to the Government

* The Peer Review Summary Report will not be submitted to, reviewed, or approved by the Contractor.

Applicable Performance Standards
The acceptance of the contract deliverables shall be based on three performance standards: (1) The reports shall be completed in accordance with the required formatting and content in **Annex 1**; (2) The reports shall address each ToR as specified **Annex 2**; and (3) The reports shall be delivered as specified in the schedule of milestones and deliverables.

Travel
All travel expenses shall be reimbursable in accordance with Federal Travel Regulations (Travel resources | GSA), and all contractor travel must be approved by the COR prior to the actual travel. Any travel conducted prior to the receipt of proper written authorization from the COR

will be done at the Contractor's own risk and expense. International travel is authorized for this contract. Travel is not to exceed $12,000.00.

Confidentiality and Data Privacy

This contract may require that services contractors have access to Privacy Information. Services contractors are responsible for maintaining the confidentiality of all subjects and materials and may be required to sign and adhere to a Non-disclosure Agreement (NDA).

NOAA Fisheries Project Contact:

Cody Szuwalski
Research fisheries biologist
AFSC/NMFS/NOAA
7600 Sand Point Way, Seattle, WA, 98115
cody.szuwalski@noaa.gov

Annex 1: Peer Review Report Requirements

1. The CIE independent report shall be prefaced with an Executive Summary providing a concise summary of the findings and recommendations as specified in the ToR.

2. The main body of the reviewer report shall consist of a Background, Description of the Individual Reviewer's Role in the Review Activities, Summary of Findings for each ToR in which the weaknesses and strengths are described, and Conclusions and Recommendations in accordance with the ToRs.

   a. Reviewer should describe in their own words the review activities completed during the panel review meeting, including providing a brief summary of findings, of the science, conclusions, and recommendations.

   b. Reviewer should discuss their independent views on each ToR even if these were consistent with those of other panelists, and especially where there were divergent views.

   c. Reviewer should elaborate on any points raised in the Summary Report that they feel might require further clarification.

   d. The CIE independent report shall be a stand-alone document for others to understand the weaknesses and strengths of the science reviewed, regardless of whether or not they read the panel summary report. The CIE independent report shall be an independent peer review of each ToRs and shall not simply repeat the contents of the summary report.

3. The reviewer report shall include the following appendices:

   Appendix 1: Bibliography of materials provided for review
   Appendix 2: A copy of the CIE Performance Work Statement
   Appendix 3: Panel Membership or other pertinent information from the panel review meeting.

Annex 2: Terms of Reference for the Peer Review

Review of Eastern Bering Sea Tanner Crab Assessment

CIE reviewers are contracted to complete their independent peer review based on the ToRs. Therefore, the CIE-NOAA Fisheries review and approval process is based on whether the CIE independent reports addressed each of the ToRs.

1. Generalized Model for Assessing Crustaceans Stocks (GMACS)
    a. Comment on the pros and cons of transitioning from the status quo bespoke assessment to the GMACS framework.
    b. Are there any concerns with the GMACS framework that would preclude its adoption?
    c. If so, are there clear paths to addressing these concerns (and what are the paths)?
2. Side-by-side selectivity
    a. Are the methods presented for estimating the selectivity of the NMFS survey gear based on the side-by-side experiments scientifically sound?
    b. Are the resulting changes in the scale of the population when incorporating the side-by-side data into the assessment defensible and appropriate for providing management advice?
    c. If not, how would the reviewer approach the provision of management advice using assessment output that incorporates the selectivity experiment data?
3. Reference points
    a. Are F35%-based reference points appropriate for Tanner crab in light of terminal molt and the discrepancy between maturity and selection in the fishery?
    b. If not, can the reviewer provide more appropriate alternatives for reference points?
    c. Is morphometrically mature male biomass the appropriate metric for use in the harvest control rule?
4. Two area model
    a. Under what circumstances would a two-area model for Tanner crab be recommended?
    b. Are there any circumstances under which a two-area model would expressly not be recommended?
    c. Are there any data that could be gathered (or analyses performed with existing data) to inform this decision?
5. Hybrid snow/Tanner crab
    a. How should hybrid crab be treated in the assessment?

b. How should hybrid crab be treated in the development of reference points and harvest control rules?

c. Are there additional data streams that would be helpful to gather or experiments to be performed that would better inform how to treat hybrid crab in assessment and management?

Reviewers must provide details on the basis of their responses to the review questions.

Annex 3: Tentative Agenda for the eastern Bering Sea Tanner crab assessment

June 9-11, 2026

Chair: Cody Szuwalski (cody.szuwalski@noaa.gov)
(to be expanded as presentations develop)

Day	Time	Topic	Presenter
June 9	9-12	Welcome and overview	Szuwalski
	930-12	Introduction to Tanner crab assessment	Stockhausen
	1-2	Survey data description	Kodiak team
	2-3	Catch data description	ADFG
	3-5	Discussion and identification of requests	
June 10	9-12	Report back on CIE requests	Stockhausen
	1-5	Discussion and identification of final requests	
June 11	9-12	Report back on CIE requests	
	1-5	Wrap up	

Appendix 3: Panel Membership or other pertinent information from the panel review meeting.

The review proceedings were chaired by Cody Szuwalski, NOAA, and the review panel consisted of Yong Chen, Martin Cryer, and Larry Jacobson representing CIE.

The following were also present in the room or online:
William Stockhausen, NOAA
Grant Adams, NOAA
Sandra Lowe, NOAA
Melissa Haltuch, NOAA
Mike Litzow, NOAA
Jim Ianelli, NOAA
Cole Monnahon, NOAA
Tyler Jackson, ADFG
Katie Palof, ADFG