

**Gulf of Alaska  
walleye pollock  
stock assessment review**

**P.L. Cordue  
Fisheries consultant  
New Zealand**

***For The Center of Independent Experts (CIE)***

**August 2012**

## Executive summary

A review of the Gulf of Alaska pollock stock assessment was conducted by three independent CIE appointed reviewers. The review consisted of a pre-meeting review of documents, participation in a review meeting at Alaska Fisheries Science Center, Seattle, from 17–20 July 2012, and the preparation of an independent report by each reviewer.

The pollock stock in the Gulf of Alaska is assessed separately from pollock in the Bering Sea and the Aleutian Islands. The fishery was initiated by foreign vessels in the early 1970s but has been fully domestic since 1988. Catches peaked at approximately 300,000 t in 1984 and were approximately 50,000 – 100,000 t from 1999 to 2010. Ecosystem studies of the Gulf of Alaska have shown that pollock is a “hub” in the food-web, being an important food source for Steller sea lion, arrowtooth flounder, Pacific halibut, and Pacific cod.

The stock assessment uses a simple age-structured model with maximum-likelihood estimation of parameters. Data come from a range of sources: catch history (1964 – present), catch sampling by observers for length and age composition; NMFS bottom trawl surveys (1984 – present); pre-1984 trawl surveys; ADFG crab and groundfish bottom trawl surveys; and acoustic and egg surveys of Shelikof Strait during winter. The stock assessment model has essentially been unchanged since 1999.

The Gulf of Alaska pollock stock assessment contains some poor structural assumptions and uses data which have not been properly reviewed as stock assessment inputs. It is also a “precautionary assessment” in that assumptions have been made to deliberately introduce bias in order to reduce the chances of overly optimistic yield estimates. The use of precautionary *assessment* to implement precautionary *management* is not an approach that I recommend. Stock assessment should be aimed at providing a “risk-neutral” or “unbiased” assessment of the status of the stock.

My main conclusions are:

- Survey design and data collection methods for assessment inputs are generally very good.
- The stock assessment model structure and assumptions are generally not good and the potential accuracy of the assessment is hindered by the use of “precautionary” assumptions.
- Stock assessment uncertainty is badly under-estimated.
- An SPR-based fishing mortality of the order of  $F_{35\%}$  or  $F_{40\%}$ , when derived from a single-species constant-M model, may be a poor proxy for  $F_{MSY}$  for stocks where natural mortality is highly dependent on the abundance of predators. Depending on the steepness of the stock-recruitment relationship and the effectiveness of the predators when pollock abundance is low, the use of such fishing mortalities could result in average levels of spawning biomass below what is generally considered desirable or safe.

My main recommendations are:

- Each time series of data should be carefully reviewed and documented. There needs to be careful consideration of the strengths and weaknesses of each survey in terms of comparability across the whole time series (i.e., changes in design, gear, timing, and protocols).

- The acoustic surveys should be focused on estimating spawning biomass and not total biomass. To the extent possible and/or necessary, existing data should be reanalyzed to produce the best estimates of spawning biomass and abundance indices for 1 and 2 year olds.
- More trawling (and different types of trawling) needs to be done during acoustic surveys to better establish the length-strata and to determine the true length and age composition of pollock marks.
- An acoustic survey of all major spawning grounds during winter in the same year should be considered if feasible.
- The existing stock assessment model needs to be restructured.
  - For the base model(s) use only the best quality data with defensible assumptions.
  - The NMFS trawl  $q$  should be estimated and an informed prior should be used.
  - A Shelikof Strait spawning-biomass time series should be used (for which selectivity can be assumed to equal maturity).
- A new stock assessment model which incorporates trends in predator abundance should also be developed.
- Given the importance of pollock in the Gulf of Alaska ecosystem the conservative approach to management which is currently adopted does seem advisable. Therefore, it is important that when the move is made to a “risk-neutral” stock assessment that there is a simultaneous move to a more conservative control rule.

## **Background**

Periodically the North Pacific Fishery Council (NPFCC) has its stock assessments peer reviewed. The Gulf of Alaska (GOA) pollock stock assessment was last reviewed in 2003 and was well overdue to be considered again.

I am one of three CIE reviewers who participated in the review which consisted of a pre-meeting review of documents, participation in a review meeting at Alaska Fisheries Science Center, Seattle, from 17–20 July 2012, and the preparation of an independent report by each reviewer. This report presents my findings and recommendations in accordance with the Terms of Reference (ToRs) for the review (Appendix 2, annex 2).

## **Review Activities**

### *Pre-meeting*

The main documents provided for the review were made available in a timely manner. After reading most of the documentation I was concerned that there was little detail on the comparability of survey estimates within each time series. From reading the 2003 CIE review reports (Godø 2003, Haddon 2003, which I found online) I realized that there were serious concerns with regard to the NMFS bottom trawl survey. Therefore, I made an email request for the assessment team to put together some more information on the comparability of the surveys:

It would be useful to see, for each series which has an assumed single  $q$ , a table/graph showing any changes which have occurred in timing, protocol, or gear for that series (e.g., a trend in start dates for a trawl survey or a change to the standard tow duration).

I was also concerned that the acoustic surveys were being used to produce proportion-at-age estimates when there appeared to be no technically fully-defensible method for producing such estimates. I requested some clarification of the methods:

One of the issues with the acoustic time series is how the design achieves sensible estimates of length or age structure. There appears to be a weighting issue if there is any (length/age) variability between marks within a length-stratum. It would be useful if someone could show, for example using the most recent survey, the spatial location of the length strata, how they were determined, how many tows were done within each stratum, how many marks and mark types were in each stratum, and the length frequencies from individual tows (as well as the combined length frequency for each length-stratum). It would be great if they could also explain why the method used for combining individual length frequencies, within a length stratum, is appropriate.

The purpose of these requests was to give the assessment team some warning of these issues, rather than surprising them with requests for more information at the meeting.

### ***During the meeting***

The first two days followed the Agenda (Appendix 3) closely with excellent presentations given by the presenters. My pre-meeting requests had been addressed by the team and I believe that extra slides had been added to presentations to help clarify the issues.

There was considerable discussion with regard to proportion-at-age data from the acoustic surveys. My contention was that there was no technically defensible method to produce unbiased estimates of proportion-at-age. This point was never conceded by the acoustics team. Their contention was that they had used the best method available. I didn't actually disagree with that point. I tried to explain that if one were to attempt to provide estimates of proportion-at-age from the acoustic surveys then what they had done was pretty good. However, my point was that it couldn't actually be done in a fully defensible way (because of the intrinsic difficulties of obtaining unbiased length/age samples of ensonified layers/marks using trawl gear). They acknowledged the mesh-selectivity issue (see Williams et al. 2011), but didn't accept my wider concerns.

There was also discussion of the " $q = 1$ " issue for the NMFS trawl survey. The reviewers were unanimous in believing that this was an inappropriate assumption (a view shared by the CIE reviewers in 2003). Dr. Dorn acknowledged that it wasn't really a suitable assumption for stock assessment but did suggest that the more important issue was "good management outcomes".

Discussion on the suitability of a single  $q$  for the NMFS trawl time series was on-going during the meeting. There was an excellent slide in the presentation on the surveys which showed, for each vessel in each survey, the occupation dates of trawl stations versus longitude. This slide made it clear that the first two surveys were dubious members of the time series just in terms of timing. The issue of non-standard trawl vessels and gear being used in these surveys was also discussed and the group agreed that it was difficult to argue for a single  $q$ .

In the final two days of the meeting a "working model" was explored by the group. An attempt was made to eliminate poor assumptions and data inputs to see if a reduced set of data could be reasonably well fitted by the model. Several runs were done, but these were all of an exploratory nature rather than an attempt to get a defensible assessment.

The main features of the working model, relative to the existing assessment, were: remove the historical (pre-1984) trawl time series; remove the 1984 and 1987 points in the NMFS trawl survey time series; fit to a spawning biomass time series from the Shelikof Strait acoustic surveys; exclude the acoustic proportion-at-age data (or, alternatively, fit to it as a time series with its own selectivity); use time-blocks for the fishery selectivity (rather than a random walk); and estimate the  $q$  for the NMFS trawl time series (or fix it at 0.5, 0.75, and 1).

The explorations were interesting and informative. When the 1984 and 1987 points in the NMFS time series were removed, the visual contradiction between the biomass indices within the model largely disappeared. However, there was still a problem fitting the large decline in the acoustic time series and the model wanted to put in some very large recruitments prior to the time of the decline in order to fit it. In the final few runs, it was apparent that the fishery catch-at-age data were in some conflict with the acoustic time series and that there was a relative-weighting issue that would need to be explored.

### ***Post-meeting***

I decided it would be useful to construct a simple model to investigate the suitability of  $F_{35\%}$  as a proxy for  $F_{MSY}$  when natural mortality was driven by predator abundance (as per ToR 4). In the absence of such an investigation I thought I would have little basis for drawing conclusions with regard to ToR 4. Dr. Dorn had mentioned at the meeting that it was really in the “too hard” basket, but they had put it in the ToRs to see what the reviewers could offer. I think my investigations did yield some useful insights (see ToR 4 in the Summary of Findings).

I also sent an email request to Dr. Dorn with regard to the exact meaning of  $B_{40\%}$  within the NPFC:

Tier 3 requires a reliable estimate of  $B_{40\%}$ . But that requires that the stock-recruitment curve is reliably estimated which means that  $B_{MSY}$  is reliably estimated (which moves the stock into tier 2). Also, if  $B_{40\%}$  is given its usual meaning then as steepness moves lower the control rules become increasingly more aggressive (as the kink comes in at lower biomass levels).

I hope you see my problem. From the tier structure it looks like  $B_{40\%}$  actually means 40%  $B_0$  – which is a notational problem – but if  $B_{40\%}$  actually means  $B_{40\%}$  then there appears to be a problem with the tiers.

Can you clarify please?

Dr. Dorn sent a helpful reply and did confirm that the NPFC uses the normal definition of  $B_{40\%}$ .

## **Summary of findings**

The current GOA pollock stock assessment has been largely unchanged since 1999. The degree of stability is not a reflection of the quality of the assessment which is poor. The data have not been adequately screened and reviewed for use in an assessment and there are a number of poor structural assumptions made in the model.

Each of the ToRs are specifically considered below.

- 1. Evaluate and provide recommendations on data collection procedures and analytical methods used to develop assessment model input.*

### ***Catch history and catch sampling***

The catch history appears to be well-defined. The fishery developed in the 1970s and catches peaked at a bit over 300,000 t in 1984. Catches from 1999 to 2010 have been approximately 50,000 – 100,000 t (Dorn et al. 2011). Changes in fleet structure are described and catches are split by fleet in stock assessment documents. It would be useful to also document the change in the seasonality of catches (i.e., tabulate and graph the full time series of catches showing the seasonal distribution). Similarly, a full description of the changes in the spatial distribution of the fishery would be helpful. Such changes are important in determining the fishery selectivity and, in particular, helping to decide which time-blocks (for fishery selectivity) should be used in the stock assessment.

Catch sampling by observers seems adequate. Numbers of fish measured (generally > 10,000) and aged (approximately 2000 since 1999) seem adequate, although numbers were a bit low during the mid 1990s (especially 1994, see Table 1.6 in Dorn et al. 2011). The use of two age-length keys should be adequate to cover the growth of fish during the year, particularly as the fleet captures few 1 year olds.

Ageing methods were hardly discussed at the review, but I noted the comments of Haddon (2003) who was concerned that the length of a fish was used by the age-reader to help in age determination. He was concerned that this compromised the assumed independence of the age-length key and the length frequency. I don't share this concern. The use of general knowledge with regard to the age of fish of a certain length (from well-defined length modes) is not introducing a statistical dependence – the sampling error associated with the length frequency is not being introduced into the age-length key.

### ***Egg surveys***

The egg surveys to estimate spawning biomass in Shelikof Strait in 1981 and 1985–1992 were not discussed during the review. I note that the 2003 CIE review also seems not to have discussed this time series. I assume that the time series has been reviewed previously to ensure that the surveys in each year are comparable in terms of methods used and survey timing (Dorn et al. 2011, notes that the 1981 estimate is questionable – perhaps then it should not be used). Documentation defending these data as a valid spawning biomass *time series* should be produced and made available for future stock assessment reviews.

### ***Trawl surveys***

An analysis of pre-1984 bottom trawl survey data is given in Dorn et al. (2011) but there is no reference to another document which contains full details of the analysis. I assume that this work has not been fully documented. The analysis provides support for the hypothesis that pollock abundance in GOA was very low from 1961 to 1971 (prior to the start of commercial pollock fishing) and that there was an enormous increase in abundance in the mid 1970s (Table 1.12, Dorn et al. 2011). This is important context for the pollock stock and the work should probably be fully documented. As a stock assessment input it is hard to defend because of the untenable assumptions of constant area effects over approximately 30 years; and the large scaling factor of 3.84 used to link this “time series” with the NMFS trawl survey.

The ADFG trawl survey was only briefly discussed at the review. Initially no documents were provided on the survey but during and after the meeting a couple of reports were provided. The survey is a fixed-station bottom-trawl survey aimed at crabs and groundfish. It has been conducted most years since 1989. The survey is notable in that it captures mainly large pollock (typically 40–70 cm, see Fig 1.11, Dorn et al. 2011). A potential weakness of the survey, for pollock, is that it covers a very limited area in comparison to the NMFS bottom trawl survey. It is not clear to what extent it covers the areal distribution of pollock. It would be worthwhile to analyze the NMFS survey to determine what proportion of pollock biomass (by size class) is in the area covered by the ADFG survey. This would give some idea of the areal availability of pollock in the ADFG survey. As with the other stock assessment time series, a document analyzing the full time series in terms of comparability over time of survey methods, timing, and protocols is needed.

The NMFS bottom trawl surveys are random stratified multi-vessel trawl surveys that were triennial from 1984 to 1999 and biennial since then. However, the time series is not consistent over the whole time frame. It really isn't until 1996 that consistent protocols and vessels were

used. In 1984 and 1987, non-standard Japanese vessels and gear were used in the surveys (Munro & Hoff 1995). Also, the stations were occupied in a different order compared to surveys from 1996 onwards and the surveys finished later (some stations were occupied a month earlier and some stations up to three months later – see GOA bottom trawl survey presentation by Dr. Martin). There was also a timing issue in 1990 and 1993 but it was not so severe. Finally, trawl duration was 30 minutes prior to 1996 and 15 minutes thereafter.

I was aware from the 2003 CIE review reports that there were issues with this bottom trawl time series. However, there was no indication in the 2011 stock assessment report (Dorn et al. 2011) that there were any issues relating to the comparability of the surveys. The timing issue in 1984 and 1987 alone is enough to question whether those two surveys should be included in the time series. When the use of the different vessels and gear is acknowledged there is no question that they should be excluded from the main time series. They could perhaps be included in the stock assessment if priors (or penalties) on the ratio of trawl-survey  $q$ s (in each year) to the main-survey  $q$  could be quantified. Inclusion or exclusion of the 1990 and 1993 points in the main time series is not so clear cut. However, the timing of the surveys is up to a month later and protocols for trawl duration and the definition of time on-the-bottom were different. It would be best to exclude them from the main time series and give them a different  $q$  with a prior on the ratio of the  $q$  to the main time series.

The use of multiple vessels in the main time series (1996 onwards) is of some concern and it would be useful if an attempt was made to standardize the time series using a GLM approach (to estimate vessel/skipper effects in particular).

### *Acoustic surveys*

Acoustic surveys of Shelikof Strait have been conducted in winter during spawning time in most years since 1981. There have also been surveys in other spawning areas starting in 1994, but more consistently since 2002. There have been three attempted surveys of the GOA in summer, but for various reasons they have not covered the full spatial distribution of pollock.

The survey methods generally appear to follow best practice. Calibrations are conducted before and after the surveys; systematic transects are used with a random starting point; trawling is done with mid-water and bottom gear for target identification and to determine pollock length and age composition. The survey data are post-stratified into strata with different length-frequencies which attempts to deal with systematic changes in length structure within the survey area. The acoustics team has also conducted several experiments to investigate the validity of the assumptions underlying the estimates of biomass and age frequency. They have investigated vessel avoidance, mesh selection of the mid-water gear, and the use of multiple frequencies for target identification.

However, I do have some concerns about the analysis of the acoustic survey data. The longest time series of estimates is in Shelikof Strait which is thought to be the main spawning ground. However, in most years a large population of non-spawning fish is also present in the area. This creates a challenge for the acoustic method because the presence of multiple length modes in the surveyed population makes the estimation of the ratio of mean weight to mean backscattering cross-section very problematic (this ratio is used to scale the backscatter to biomass). The problem exists for length-strata which have multiple length modes (e.g., see the report on the 2010 survey, Guttormsen & Jones 2010). In order to accurately estimate the scaling ratio (within a length stratum), an unbiased estimate of the length frequency is required.

The current survey design uses opportunistic sampling of “dense” pollock marks with mid-water or bottom trawl gear to obtain the length frequency samples (and to determine the length-strata). This method cannot be expected to provide unbiased estimates of the proportion in each length class (e.g., length modes at ages 1, 2, 3, 4+ which appear typical of the pollock mixed age-class layers) because of non-random selection processes: mesh selectivity (e.g., age 1 fish goes through the meshes – see Williams et al. 2011); vertical availability (layers may have size-related vertical structure); vulnerability (size-related avoidance reactions may exist e.g., larger fish have a stronger dive reaction); spatial structure within length strata (e.g., shallower areas of a length stratum having more smaller fish than the deeper areas; or western areas having more larger fish than the eastern areas).

Some illustrative examples are given in Appendix 4 which show how the non-random selection processes of the sampling method can lead to biases in the estimation of proportion-at-age/length (which will flow through to bias in the estimation of the scaling ratio and hence to the biomass estimates). The examples also show that just correcting for the low selectivity of 1 year olds, due to mesh selection, will not solve the problem – it is only one of the sources of bias.

The presence of multiple length modes (or multiple species) in acoustic marks/backscatter makes total biomass estimation problematic. Whatever method is used to obtain biological samples there is a selection pattern and the relative proportions of length classes (or species) cannot be accurately estimated unless the selection pattern is known.

All of that said, the acoustic survey data are potentially still very informative as a time series. The scale of the decrease in biomass from 1981 to the present is far larger than any potential systematic trend in the bias. Indeed, one might expect that the annual biases have little systematic trend in them and that as “random” annual biases they simply feed into increased imprecision in the biomass estimates. The proportion-at-age data are likely to be useful for the younger ages, but less so for the older ages. The contrast between strong and weak year classes, as 1 and 2 year olds, should be strong enough to allow signals in the estimates to overcome the biases. As with the biomass, with perhaps “random” annual biases, it may be that the data are adequately modeled with low effective sample sizes.

The acoustic surveys are currently analyzed to produce total biomass estimates and proportion-at-age. This is not a good formulation for the stock assessment model. Instead they should be analyzed to produce spawning biomass estimates and 1 and 2 year old biomass indices. The main reason for this is that the (age-based) selectivity associated with the biomass estimates is very different from the selectivity associated with the proportion-at-age data. The selectivity of the proportion-at-age data is driven by the trawling gear and protocols (and biases associated with them). Whereas, the selectivity associated with the biomass estimates is mainly determined by the fish that are ensonified by the acoustic beam. Simply put, if there were equal proportions of each age class in the survey area, it would be expected that the acoustic transects, on average, would ensonify roughly equal proportions of each age class (though there could be some age-specific avoidance and under-sampling in the shadow zone). However, the trawl sampling would yield proportion-at-age estimates very different from what was present.

There are a number of ways to reanalyze the acoustic data to produce spawning biomass estimates and 1 and 2 year old biomass indices.

The quickest method would be to use the existing length-strata in each survey to produce estimates for each category (mature, 1 year old, 2 year old) for each stratum (and then sum across strata). This needs, for each stratum, the proportion-mature-at-length key and the length

frequency (no age data are needed as 1 and 2 year old fish can be identified by length modes). The *very* quick method of producing a spawning biomass estimate by applying a single proportion-mature-at-length key to the biomass at length for the whole survey should not be used because proportion-mature-at-length will be very different in the spawning-biomass length stratum compared to the others.

Another method that should be considered is to reanalyze each survey with a specific view to obtaining the best estimate of mature biomass and 1 and 2 year old biomass for each survey. This may involve a different post-stratification and in particular reconsideration of the dividing line between spawning marks and non-spawning layers. If spawning marks are just completely mixed in with non-spawning fish this may not be possible, or perhaps it will be possible in some years and not in others.

The amount of trawling in recent years has reached very low levels (less than 20 trawls each year since 2006). This seems hardly enough to properly define the length strata let alone get any idea of the vertical and spatial structure of the acoustic marks (i.e., how length structure varies). Apparently there are data from earlier years collected to investigate the structure within layers – these data should be fully analyzed and documented. Also, further efforts should be made during surveys to ascertain the true length-structure of important layers. For example, a bottom referenced layer extending 120 m off the bottom should not just be fished by dipping the mid-water net into the top the layer – large fish may be diving and avoiding the net. The camera system with the open cod-end could be used, but conventional nets can also be tried. The mid-water gear can be progressively lowered through the layer (chasing the fish down) until it is fishing near the bottom; it could also be started near the bottom and hauled up through the layer. Off course, the layer could also be fished with bottom gear.

The use of a length-target strength relationship from 1996 is a worry (Traynor 1996). There must have been more recent data collected and during winter rather than summer. Also, it shouldn't be assumed that the relationship is necessarily linear over the full length range. It may not be, depending on how well-developed the swim-bladder is for young fish (e.g., do 1 year olds have an inflated swimbladder?). Certainly, a slope of 20 should not be forced in the relationship (Mclatchie et al. 1996).

The timing of each survey should be checked to make sure that each year is a valid entry in the time series. The spawning biomass is apparently very low in some years and it is important to know whether this is a timing issue or whether the fish simply didn't spawn in Shelikof Strait in those years. Recent attempts to cover other spawning grounds in the same year is a good idea. Spawning biomass estimates for the other grounds (some of which are also contaminated by immature fish) should also be produced, and checked for timing, so that spawning biomass can be compared across areas (much more useful than total biomass).

I am not sure if it is possible to cover all of the spawning grounds in winter at the appropriate time or not (with one vessel). Certainly, it would be very useful to have a total spawning biomass estimate for the whole stock in at least one year. Attempts to obtain a stock-wide estimate with a summer acoustic survey have been unsuccessful so far. In theory, that would also be very useful but I am not sure that it is feasible given the problems with target identification, mixed species layers, and multiple-length mode pollock marks. If spawning grounds and timing are readily identifiable, without too many immature fish contaminating the spawning marks, then the use of extra resources in winter to obtain a total spawning biomass may be a better option.

I also have some minor concerns which could be addressed in a document which examines the comparability of the whole Shelikof Strait time series (and the surveys of other spawning areas). The quantity of eulachon caught in the Shelikof Strait surveys in some years is a worry in that the vulnerability to the trawl gear may be very low. The target strength is very low, because they do not have a swimbladder, but if the vulnerability to the trawl net is also very low then they may still contribute significantly to the backscatter – this concern should be put to rest by analyzing the eulachon data from the pocket net experiment (Williams et al. 2011) and doing some calculations. Also, the use of a hull-mounted transducer (albeit on a centerboard) means that there will be signal attenuation due to the wind-induced bubble layer (which, when present, will be deeper than the transducer). Vessel motion will also cause some signal loss (i.e., pinging while the transducer is pointing in one direction and listening when it is pointing in another). Corrections for the bubble attenuation and vessel motion should be calculated for each year (assuming that pitch and roll data are collected during the surveys). These corrections will not be large, but they will vary from year to year depending on prevailing weather conditions.

2. *Evaluate and provide recommendations on model structure, assumptions, and estimation procedures.*

A simple population dynamics model is fitted to the various data inputs using maximum likelihood/Bayesian methods implemented in AD Model Builder (ADMB). There are some penalty functions as well to keep parameters in sensible space when there is little data to constrain them or where there is prior information. The general approach is standard. However, the current assessment is very poor in terms of providing a reliable stock assessment.

The treatment of individual data sets deviates from good practice in several instances. The most egregious fault is something which appears to be an historical relic where the trawl-survey  $q$  for the NMFS bottom-trawl survey is assumed to be equal to 1. In Dorn et al. (2011) this is justified with the statement: “The NMFS bottom trawl survey catchability was fixed at one in this and previous assessments as a precautionary constraint on the total biomass estimated by the model”.

The use of “precautionary *assessment*” to implement the precautionary principle in fisheries *management* is a serious mistake. The purpose of stock assessment should be to obtain a “risk neutral” assessment of the status of the stock (its absolute biomass level and its level relative to reference points). Any necessary assumptions should be based on a “best guess” and an attempt to avoid the introduction any large positive or negative biases in the estimation procedure. A correct implementation of the precautionary principle in fisheries is through fisheries management where precautionary reference points and controls rules can be used.

***NMFS trawl***

The current assessment uses the full times series from 1984 with a single  $q$ . As explained under ToR 1, a single  $q$  is inappropriate for this time series because of vessel, gear, timing, and protocol issues. The time series should be started in 1996. Earlier surveys could be included with different  $qs$  if suitable priors can be developed on the ratios of the  $qs$  to the main-series  $q$ . The age-frequency data could perhaps be retained for the whole time series with the same selectivity, but strictly speaking different selectivities would also be needed.

The  $q$  for the time series should be estimated in the model with an informed prior. During the review meeting, upper and lower bounds for the three components of the trawl  $q$  were discussed: areal availability (0.8, 1); vertical availability (0.7, 1); and vulnerability (0.8, 1). These

assumptions give overall bounds of (0.45, 1). Clearly, the assumption of  $q = 1$  for this time series is not a good choice.

#### ***ADFG trawl***

This time series needs to be reviewed to see if the use of a single  $q$  over the whole time series is justified. There also needs to be a check that the areal availability is not so low that the reliability of the time series as an abundance index would be brought into question. Ideally, an informed prior for the  $q$  would be developed.

#### ***Egg surveys***

This time series needs to be reviewed to see if the use of a single  $q$  over the whole time series is justified.

#### ***Historical trawl survey time series***

This time series does not add anything useful to the assessment. It should be used in a qualitative sense (providing some context for the stock dynamics – highly variable abundance) but is not quantitatively robust enough to be used in a stock assessment.

#### ***Acoustic surveys***

The Shelikof Strait time series is currently fitted as a total biomass index and proportion-at-age. The selectivity estimated within the model for the acoustic data implies that Shelikof Strait, during winter, is a primary area for juveniles and, relatively speaking, contains little spawning biomass. This is contrary to common sense and appears to be an artifact of inappropriate assumptions made in the model, particularly with regard to the acoustic data.

As explained in TOR 1, the selectivity associated with the proportion-at-age data should be expected to be very different to the selectivity for the acoustic biomass index. The difficulty arises in the model that if total biomass is fitted then the only selectivity available to use is that estimated from the proportion-at-age data. However, a solution to this problem is to fit a spawning biomass index from the acoustic surveys; the selectivity is already prescribed, simply being the maturity ogive.

The proportion-at-age data could still be fitted in the model, but it must be done without linking it to the acoustic biomass (it goes in as an index of proportion-at-age in winter with its own selectivity). However, I think it would be preferable to rework the acoustic data to provide 1 year old and 2 year old biomass indices (which are where the main information lies) and to only fit the proportion-at-age data in sensitivity runs.

For the spawning biomass time series there would be no need to put a break in the time series for the changeover from the Biosonics to EK500 systems due to different thresholds. It should be safe to assume that most of the spawning biomass is coming from dense marks which would not be affected by thresholding (this could be checked for some of the surveys). I am also doubtful that a correction should be made for the change to the quieter vessel – I would expect that spawning fish are deeper and less susceptible to noise disturbance than non-spawning mid-water fish (this could perhaps be checked by examining the depth-dependence of avoidance in the existing experimental data).

### ***Maturity at age***

In the current model, maturity at age is taken to be the average maturity ogive from the acoustic-survey estimates in Shelikof Strait. The annual estimates are highly variable which is hardly surprising given this is the wrong place and the wrong time of year to estimate maturity-at-age within the population. Shelikof Strait during winter has a spawning population (which presumably migrated into the area for spawning) overlaid on a population of immature fish (possibly resident). It would only make sense to estimate maturity-at-age in Shelikof Strait if there were equal proportions of the mature and immature populations present during winter. It seems likely that a greater proportion of the mature fish will be there (since it is the main spawning ground) and therefore the estimates are probably biased high (although it is not clear that the other biases associated with the acoustic proportion-at-age data would not override this bias). It might be useful to analyze the data a little bit more because, under the assumption of a constant proportion of immature fish in Shelikof, the data could contain information on the proportion of the spawning population in Shelikof each year (and hence to the potential variability in the proportion spawning in Shelikof).

### ***Fishery catch-at-age***

In the current model, the fishing selectivities are allowed to vary annually with parameters following a constrained random walk. However, it is clear from the model fits that the constraints are not strong enough to stop the model essentially fitting noise (see Fig 1.17 in Dorn et al. 2011; in 1991 the fleet apparently managed to successfully target 9 year old fish). It would be more parsimonious to examine the catch history in terms of fleet composition, spatial and seasonal distribution of the catch to determine appropriate time blocks within which to fit constant selectivities.

### ***Initialization***

In the current model a constrained initial age structure is estimated in 1961. There are three main options for initial conditions that should all be tried: biomass and age-structure in equilibrium at  $B_0$ ; age-structure in equilibrium at  $B_{initial}$  (not necessarily equal to  $B_0$ ); and a freely estimated initial age structure. AIC could perhaps be used to decide which is best. Priors are likely to be needed for the last two options.

### ***Stock recruitment relationship***

This is where some precautionary principles could legitimately be applied as the choice of stock recruitment relationship influences MSY-based reference points, SPR-based biomass levels (e.g.,  $B_{40\%}$ ) and will have an effect on medium and long term projections. It should not have too much effect on estimated stock status (unless there are heavy penalties to force recruitment to follow the stock-recruitment curve).

I was surprised to see the assumption of no stock-recruitment relationship, which is equivalent to a Beverton-Holt stock-recruitment relationship with steepness equal to 1 (i.e., the same expected recruitment at all levels of spawning biomass).

I would recommend using a Beverton-Holt stock recruitment relationship with a best-guess at steepness (erring on the low side if in doubt). An assumption is needed for the stock assessment and assuming no relationship is unlikely to be true. The apparent “explosion” of biomass in the

mid 1970s, when biomass prior to then was at low levels, certainly suggests high recruitment variability, but does not guarantee a high value for steepness.

### ***Age range***

Currently only ages 2-10+ are modeled in the population. It would be more natural to model a wider range and certainly to include age 1 since there are data available on them. Also, if  $M$  is estimated in some future runs it may help to have the data on older ages included as individual age classes. Moving to a wider age range shouldn't have too much impact on results – if it did it would need to be investigated (and might be instructive).

### ***Natural mortality***

This is fixed at 0.3 at all ages which may be a bit low given the results from Hollowed et al. (2000) and that domed selectivities are estimated in the model for the fishery. Dorn et al. (2011) mention a recommendation that a low  $M$  assumed in the model is precautionary. I have already noted that precautionary assessment is a bad idea. Other options for  $M$  should be explored including attempting to estimate it in the model after fixing fishery selectivity to be flat-topped in one of the time blocks. The use of a maximum age approach, for estimating  $M$ , is very problematic if pollock abundance is driven by predator abundance (as there will be periods of time when  $M$  is low and some fish will reach an old age).

### ***Likelihoods***

The likelihoods used are fairly standard, but are technically wrong in the case of fishery independent surveys which are designed to provide unbiased biomass estimates. These should be modeled as mean-unbiased lognormal random variables:  $X_i = qB_i\varepsilon_i$  where  $E(\varepsilon_i) = 1$ . This assumption gives a slightly different likelihood than that used but it will make little difference except when the CVs of the indices are quite large.

### ***Data weighting***

This is the “big” issue for many assessments, particularly if there are contradictory signals from different data sets in the model. In the case where different abundance indices are essentially contradictory the solution is easy – they should not be included in the same model run because the model will “average” the signals – thus the “truth”, which may lie with one or the other, will not be found (Schnute & Hilborn 1993). When the signal in a biomass time series is over-ridden by composition data, the best solution is usually to down-weight the composition data (Francis 2012). The point being that if the biomass time series is valid, then it is far more likely to contain valid information on the trend in abundance than composition data (where abundance signals are usually confounded with uncertain selectivity).

In the review meeting, the exploratory runs that were done with a “working model” suggested that the fishery catch-at-age data were given too much weight with effective sample sizes of the order of 300. A likelihood profile over mean recruitment showed that the catch-at-age favored lower levels of biomass compared to the acoustics time series. At effective sample sizes of 300, the catch-at-age was moving the biomass estimates to much lower levels than the biomass indices indicated. That said, with less weight on the catch-at-age data, it appeared that the model produced some very high estimates for some cohorts which were not supported by the age composition (but were needed to better fit the abundance indices). There is no easy answer for

this. It will be a matter of trying different weighting schemes, and constraints on recruitment deviations, to eventually arrive at a base model and a good number of sensitivity runs.

### ***Future stock assessment models***

It may or may not be possible to find a satisfactory fit to the pollock data in a single-species model which does not account for variable natural mortality due to changes in predator abundance. Nevertheless, the first step should be to make the attempt. However, given the evidence from ecosystem models that pollock are an important hub in the GOA food-web and that predator numbers have changed substantially over the last few decades it seems sensible to (eventually) move to a new, more complicated stock assessment model. Initially, any such models would simply be a sensitivity test for the base constant-M model(s).

I would not recommend a full multi-species model for stock assessment as too many assumptions are required and too little data are available. In particular, I think that absolute abundance indices of predators must be treated cautiously as must estimates of absolute consumption rates. The approach taken by Hollowed et al. (2000) of modeling predation as a number of “fisheries” in a simple age-structured model is appealing except that it did rely on absolute abundance estimates for the predators and the consumption rates. However, conceptually, something at that sort of level of complexity seems appropriate.

The simplest approach is to follow Hollowed et al. (2000) by modeling predator mortality as fisheries. However, rather than assuming predator numbers and consumption rates are known, a grid of assessment runs should be presented which cover the range of plausible predator numbers and consumption rates.

An alternative is to modify a single-species model to include additional natural mortality from specified predators and use estimates of relative predator abundance to drive the additional natural mortality. An example of the type of model that could be used is given in Appendix 5, although the equations would need to be generalized to an age-structured model. The model has a total predation rate  $P$  which is the sum of the predation rates,  $P_i$ , from individual predator species. Generalizing the notation to include years indexed by  $y$ , each predation rate can be parameterized as a predator-number *index* ( $N_{i,y}$ ) multiplied by a consumption rate ( $c_i$ ) and a scaling coefficient ( $q_i$ ):

$$P_{i,y} = q_i N_{i,y} c_i$$

The point of this parameterization is that absolute abundance and consumption rates are not needed as the  $q_i$  can be estimated in the stock assessment model. It seems likely that some priors will be needed on the  $q_i$  but these should be easy to develop given there are estimates of absolute predator number and consumption rates (so the prior for each  $q_i$  could be centered on 1). The pollock-length stomach data from the different predators, rather than being fitted in the model should be used externally to specify plausible age-based selection for each predator – and some predators may need to be split into different classes (e.g., split into juvenile and adult).

3. *Evaluate and provide recommendations for the reporting of assessment results and characterization of uncertainty.*

Confidence intervals, for a single base model, are provided in the current assessment for recruitment and spawning biomass (Dorn et al. 2011). I assume the intervals come from an approximation using a likelihood profile (but I didn't notice any reference to the method in the document). Projections are done by using the terminal age distribution from the point estimate and resampling recruitments in some way (again, I am not sure of the exact method – it was briefly described in the Results section of Dorn et al. 2011).

The current characterization of uncertainty is poor. One problem is that uncertainty is severely under-estimated because of the assumption of absolute trawl biomass indices ( $q = 1$ ) and high effective sample sizes for the fishery catch-at-age data. Another problem is that alternative scenarios have not been fully explored and/or are not presented in the assessment document. It would be useful to present a *summary* of a fairly standard set of sensitivity runs each year (e.g.,  $M$  higher and lower and/or estimated and/or age specific, steepness higher and lower, different initialization assumptions). When predator-abundance driven models are developed, they should also be used to produce a range of runs, initially just as sensitivities to the constant- $M$  base model(s), but perhaps eventually providing a base model.

I suspect that most of the uncertainty will be in alternative model assumptions rather than being driven by observation error. However, given that priors will probably be needed it would make sense to take some of the runs through to full MCMCs to produce posterior distributions. The full Bayesian method also provides a good framework for projections since parameter uncertainty as well as “future” (stochastic) uncertainty is automatically incorporated.

4. *Evaluate and provide recommendations on  $F_{35\%}$  spawning biomass per recruit as an appropriate proxy for  $F_{MSY}$  under non-stationarity in vital rates. Also evaluate and provide recommendations on the  $B_{35\%}$  biomass reference point as a proxy for  $B_{MSY}$ .*

There has been a lot of Management Strategy Evaluation (MSE) work on GOA pollock looking at the robustness of the existing control rule (A'Mar et al. 2008, 2009a, 2009b, 2010). It was suggested during the review meeting that the reviewers look to the MSE work in regard to this ToR. However, I am unwilling to draw any conclusions about the suitability of  $F_{35\%}$  from the work. The work was primarily concerned with the performance of the control rule as it would operate with the existing stock assessment. The existing stock assessment is badly flawed. It assumes that certain data are available with low CVs and high effective sample sizes. The operating model uses a similar structure (in most cases) and makes the same data assumptions. Also, the simulations were neither focused on the long-term performance of the control rule nor on short term projections, but were a mixture of both. The MSE work appears to have been premature – it must at least wait for a much improved stock assessment.

In order to “evaluate”, as required in the TOR, I put together a simple single-species, single-fishery, model with a base natural mortality and additional natural mortality driven by specified numbers of predators (a “predator” model – see Appendix 5). I used the model to look at the long-term performance of  $F_{35\%}$  (as derived from a single-species model with constant natural mortality equivalent to the average total natural mortality in the predator models) when applied to predator models with different levels of steepness and predator “effectiveness”.

The “effectiveness” of the predators refers to the extent that they can continue to consume as much prey as they like when prey numbers are reduced. In this model, effectiveness was defined by the single parameter  $N_c$  which is a given percentage of the virgin prey numbers. Two values of  $N_c$  were used: 10%  $N_{virg}$  which represents very effective predators (able to maintain contact with the prey and consume them as needed even when prey numbers drop to 10% of the average virgin level); and 50%  $N_{virg}$  which represents much less effective predators (when prey numbers drop below 50% of virgin, the consumption rate of the predators is constrained and becomes proportional to prey numbers – just like  $M$  or  $F$  in the standard Baranov catch equation).

I will give a brief description of the results because they demonstrate a very important difference between constant-M single-species models and variable-M single-species models.

In the virgin state, prey numbers in the predator models ranged from 70–130 % of the average virgin level due to the specified variation in predator numbers/consumption. The annual “exploitation rate” for the predators ranged from 5–22 % whereas the base natural mortality ( $M = 0.2$ ) had an annual “exploitation rate” of 16–18 % (exploitation rate or removal rate just being the percentage that died due to the particular mortality source).

For the models with highly effective predators,  $F_{MSY}$  was quite low, being of the order of the base  $M$ , and was sensitive to steepness as expected (Table 1). Under exploitation from fishing, even at  $F_{MSY}$ , the models showed marked variability in the annual numbers of the prey population and a corresponding high variability in the exploitation rate of the predators (Table 1). For the models with the much less effective predators  $F_{MSY}$  was much higher and was still sensitive to steepness (Table 2). The variability in annual prey numbers and predator exploitation rates was greatly reduced because fishing at  $F_{MSY}$  reduced the average biomass to levels typically less than 20% of virgin (Table 2).

**Table 1:  $F_{MSY}$  and associated statistics for a predator model with highly effective predators ( $N_c = 10\%N_{virg}$ ).  $U$  denotes a removal rate (number removed divided by number available) and the subscript specifies the source (e.g.,  $U_P$  is the removal rate for the predator mortality  $P$ ).**

Steepness	$F_{MSY}$	$N_{MSY}$ (% $N_{virg}$ )	MSY (% $N_{virg}$ )	$N$ range (% $N_{virg}$ )	$U_F$ range (%)	$U_P$ range (%)	$U_M$ range (%)
0.7	0.11	42	6.5	19–66	7.3–9	8.6–43	13–16
0.8	0.18	32	8.9	11–54	11–14	9.4–52	12–16
0.9	0.29	24	12	6.9–42	16–22	10–57	11–15

**Table 2:  $F_{MSY}$  and associated statistics for a predator model with inefficient and/or non-specialized predators ( $N_c = 50\%N_{virg}$ ).  $U$  denotes a removal rate (number removed divided by number available) and the subscript specifies the source (e.g.,  $U_P$  is the removal rate for the predator mortality  $P$ ).**

Steepness	$F_{MSY}$	$N_{MSY}$ (% $N_{virg}$ )	MSY (% $N_{virg}$ )	$N$ range (% $N_{virg}$ )	$U_F$ range (%)	$U_P$ range (%)	$U_M$ range (%)
0.7	0.44	16	12	11–22	26–30	10–31	12–14
0.8	0.62	13	16	9.9–18	34–40	9.5–29	11–13
0.9	0.94	9.7	21	7.3–13	46–53	8.4–26	9.7–11

The value of  $F_{35\%}$  (0.43) from the “equivalent” constant- $M$  model ( $M = 0.344$ ) was much higher than  $F_{MSY}$  for the model with effective predators (Table 1) and lower than  $F_{MSY}$  for the model with less-effective predators (Table 2). Although a fishing mortality rate of  $F_{35\%}$  looked quite safe in the constant- $M$  model, it drove numbers to very low levels in the effective-predator model, and to approximately 20% of virgin for the less-effective-predator model (Table 3).

**Table 3: The value of  $N_{35\%}$  (% $N_{virg}$ ) for the two predator models and the “equivalent” model with a constant  $M$ . The model numbers are reduced to an average of  $N_{35\%}$  when fishing at a rate of  $F_{35\%}$  (where  $F_{35\%} = 0.43$  as derived from the constant  $M$  model).**

Steepness	Constant $M$	$N_c = 10\%N_{virg}$	$N_c = 50\%N_{virg}$
0.7	27	0.0	16
0.8	31	0.1	20
0.9	33	9.7	24

The results are fairly easy to understand. Fishing the prey population will reduce the average numbers in the population. If predators are very effective, then even at quite low stock sizes they can continue to take as much as they need – they have an advantage over other sources of mortality; high  $F$  will simply lead to very low population sizes. However, if the predators are not effective then the reduction in average biomass reduces the influence of the predators – they simply become a “normal” source of mortality and are constrained in what they can consume. The fishery can then take advantage of the high mean recruitment (in what is a stock with a high  $M$ ) and become the dominant “predator” (hence the high values for  $F_{MSY}$  in the non-effective predator model, see Table 2).

I found these results very useful in demonstrating a fundamental difference between constant- $M$  and variable- $M$  population dynamics, which, although obvious in hindsight, is nonetheless important.

In deterministic constant- $M$  models, the crucial determinant of whether  $F_{35\%}$  (or some other  $F$ ) is “safe” and/or a reasonable proxy for  $F_{MSY}$  is the slope of the stock-recruitment relationship near

the origin (e.g., Clark 2002). In variable-M models, the stock-recruitment relationship is still important, but so is the effectiveness of the predators as prey numbers are reduced. The danger of exploiting a prey population is that steepness may be lower than thought and predator effectiveness may be higher than thought. Certainly, it is unsafe to assume that an  $F_{35\%}$  derived from a constant-M model will be safe when applied to what is in reality a variable-M population.

Of course, these results also show that  $F_{MSY}$  is not necessarily a desirable fishing mortality rate for variable-M models. This is also true for constant-M models where  $F_{MSY}$  when calculated deterministically can often be close to or lower than  $20\%B_0$ .

The model is very simplistic and not tailored to pollock. It could be generalized to an age structured model and used as a stock assessment model for pollock (see ToR 2). Such a model could then be used to explore the suitability of various levels of  $F$  specifically for GOA pollock. The importance of steepness and predator effectiveness will still hold in more complex models. Therefore, I am willing to conclude that  $F_{35\%}$  (when derived from a constant-M model) could be too aggressive for pollock given that population numbers may be driven by predator abundance.

##### 5. *Recommendations for further improvements.*

The review process used by the NPFC of having strategic reviews of assessments, methods, and control rules periodically is a good approach. In terms of stock assessment review, I think it is more effective, in improving stock assessments, than having reject-or-accept reviews of assessments (e.g., SARC) or very intense but somewhat brief “workshops” which review (but often modify) an assessment (e.g., STAR Panels). However, to get the most out of stock assessment reviews, there are some improvements that could be made.

For each time series of data that are produced as an input to a stock assessment there should be existing documentation which reviews the data as a potential stock assessment *time series*. The potential limitations of the time series need to be transparent for the reviewers. This can only be the case if someone has taken the time to set out all of the strengths and weaknesses of the time series in a document that is made available to the reviewers. In this review we were provided with examples of survey reports, which just documented a *single* survey. The stock assessment report (Dorn et al. 2011) did briefly cover some of the issues with some of the time series (e.g., the reasons for three different  $q$ s for the acoustic time series) but, for example, failed to mention the problems with the NMFS bottom trawl time series.

Also, after a review, a formal document responding to the recommendations of the reviewers should be produced (this didn't happen after the 2003 CIE pollock review). This would be very helpful for future reviewers who could then see which recommendations were followed and which were not and, most importantly, the reasons why. It is also important that a formal written response is produced so that reasons for following or not following recommendations are fully thought through and recorded.

In the interests of producing the “best available information” the pollock assessment needs to be moved away from a “precautionary assessment” to a risk-neutral assessment. There also appears to be a need to adopt more conservative reference points for pollock given its apparent importance to the GOA ecosystem. It is important that both of these changes be made at the same time. It would be an error to produce a risk-neutral stock assessment and apply the current control rule. Equally, it would be over-cautious to maintain a “precautionary assessment” while moving to a more conservative control rule.

6. *Brief description on panel review proceedings highlighting pertinent discussions, issues, effectiveness, and recommendations.*

This is already covered under “Review activities” above.

## **Conclusions and Recommendations**

The existing GOA pollock stock assessment model contains some poor structural assumptions and uses data which have not been properly reviewed as stock assessment inputs. It is also a “precautionary assessment” in that assumptions have been made to deliberately introduce bias to reduce the chances of overly optimistic yield estimates. The use of precautionary *assessment* to implement precautionary *management* is not an approach that I recommend. The purpose of stock assessment should be to obtain a “risk-neutral” or “unbiased” assessment of the status of the stock. Informed by the “best possible information” managers can then take appropriate management actions.

My main conclusions are:

- Survey design and data collection methods for assessment inputs are generally very good but some improvements could be made.
- The stock assessment model structure and assumptions are generally not good and the potential accuracy of the assessment is hindered by the use of “precautionary” assumptions.
- Stock assessment uncertainty is badly under-estimated and very little sensitivity analysis is being conducted and/or documented.
- An SPR-based fishing mortality of the order of  $F_{35\%}$  or  $F_{40\%}$ , when derived from a single-species constant-M model, may be a poor proxy for  $F_{MSY}$  for stocks where natural mortality is highly dependent on the abundance of predators. Depending on the steepness of the stock-recruitment relationship and the effectiveness of the predators when pollock abundance is low, the use of such fishing mortalities could result in average levels of spawning biomass below what is generally considered desirable or safe.

My main recommendations are:

- Each time series of data should be carefully reviewed and documented. There needs to be careful consideration of the strengths and weaknesses of each survey in terms of comparability across the whole time series (i.e., changes in design, gear, timing, and protocols).
- The acoustic surveys should be focused on estimating spawning biomass and not total biomass. To the extent possible and/or necessary, existing data should be reanalyzed to produce the best estimates of spawning biomass and abundance indices for 1 and 2 year olds.
- More trawling (and different types of trawling) needs to be done during acoustic surveys to better establish the length-strata and to determine the true length and age composition of pollock marks.
- An acoustic survey of all major spawning grounds during winter in the same year should be considered if feasible.

- The existing stock assessment model needs to be restructured.
  - For the base model(s) use only the best quality data with defensible assumptions.
  - The NMFS trawl  $q$  should be estimated and an informed prior should be used.
  - A Shelikof Strait spawning-biomass time series should be used (for which selectivity can be assumed to equal maturity).
  - The relative weighting of data sets needs to be fully explored; composition data should not be allowed to dominate the signal from biomass indices.
  - A full set of sensitivity runs should be routinely performed and documented.
- A new stock assessment model which incorporates trends in predator abundance should also be developed.
  - It should be kept as simple as possible – additional complexity only added if necessary.
  - It should not rely on absolute abundance estimates of predators or absolute consumption levels
  - Use relative trends in predator abundance and informed priors on species specific parameters that scale to consumption rates.
- Given the importance of pollock in the GOA ecosystem the conservative approach to management which is currently adopted does seem advisable. Therefore, it is important that when the move is made to a “risk-neutral” stock assessment that there is a simultaneous move to a more conservative control rule.

## References

- Clark, W.G. 2002.  $F_{35\%}$  revisited ten years later. *North American Journal of Fisheries Management* 22:251–257.
- Francis, R.I.C.C. 2011. Data weighting in statistical fisheries stock assessment models. *Can. J. Fish. Aquat. Sci.* 68: 1124–1138.
- Godø, O.R. 2003. CIE Review of the Gulf of Alaska Walleye Pollock. CIE review report. 23 p.
- Haddon, M. 2003. Review of the Gulf of Alaska walleye pollock (*Theragra chalcogramma*) assessment. CIE review report. 22 p.
- Hollowed, A.B.; Ianelli, J.N.; Livingston, P. 2000. Including predation mortality in stock assessments: a case study for Gulf of Alaska pollock. *ICES Journal of Marine Science* 57:279-293.
- McClatchie, S.; Alsop, J.; Coombs, R. F. 1996. A re-evaluation of relationships between fish size, acoustic frequency, and target strength. *ICES Journal of Marine Science* 53: 780–791.
- Munro, P.T.; Hoff, R.Z. 1995. Two demersal trawl surveys in the Gulf of Alaska: Implications of survey design and methods. NOAA Technical Memorandum NMFS-AFSC-50. 139 p.
- Schnute, J.T.; Hilborn, R. 1993. Analysis of contradictory data sources in fish stock assessment. *Can. J. Fish. Aquat. Sci.* 50: 1916–1923.
- Traynor, J.J. 1996. Target-strength measurements of walleye pollock (*Theragra chalcogramma*) and Pacific whiting (*Merluccius productus*). *ICES Journal of Marine Science* 53: 253–258.

## Appendix 1: Bibliography of supplied material

The following documents were supplied before the meeting. Also, during the meeting, additional papers not listed here were supplied: 24 background acoustic papers; 2 papers on the ADFG trawl survey; the CIE reviewers reports from the 2003 review; and a paper on the 1984 and 1987 NMFS trawl surveys. Some of these additional documents are cited in my review and appear in the references.

- A'Mar, Z.T.; Punt, A.E.; Dorn et al., M.W. 2008. The management strategy evaluation approach and the fishery for walleye Pollock in the Gulf of Alaska. *In* Proceedings of the Symposium, Resiliency of gadid stocks to fishing and climate change. October 31–November 3, 2006, Anchorage, Alaska, USA. pp 317–346.
- A'mar, Z. T.; Punt, A. E.; Dorn et al., M. W. 2009a. The evaluation of two management strategies for the Gulf of Alaska walleye pollock fishery under climate change. *ICES Journal of Marine Science*, 66: 1614–1632.
- A'Mar, Z.T.; Punt, A.E.; Dorn et al., M.W. 2009b. The impact of regime shifts on the performance of management strategies for the Gulf of Alaska walleye pollock fishery. *Canadian Journal of Fisheries and Aquatic Sciences* 66: 2222–2242.
- A'Mar, Z.T.; Punt, A.E.; Dorn et al., M.W. 2010. Incorporating ecosystem forcing through predation into a management strategy evaluation for the Gulf of Alaska walleye pollock (*Theragra chalcogramma*) fishery. *Fisheries Research* 102:98–114.
- Dorn et al., Martin; Kerim Aydin, Steven Barbeaux, Michael Guttormsen, Kally Spalinger, and Wayne Palsson. 2011. Assessment of the Walleye Pollock Stock in the Gulf of Alaska. 110p. In NPFMC Gulf of Alaska SAFE.
- Dorn et al., M.W. 2004. Extending separable age-structured assessment models to evaluate trends in juvenile mortality of walleye pollock in the Gulf of Alaska. *ICES CM 2004/ FF:31*. 13 p.
- Gaichas, S.K. and R.C. Francis. 2008. Network models for ecosystem-based fishery analysis: a review of concepts and application to the Gulf of Alaska marine food web. *Canadian Journal of Fisheries and Aquatic Sciences* 65: 1965–1982.
- Gaichas, S.K.; Kerim Y. Aydin, and Robert C. Francis. 2010. Using food web model results to inform stock assessment estimates of mortality and production for ecosystem-based fisheries management *Can. J. Fish. Aquat. Sci.* 67: 1490–1506.
- Gaichas, S.K.; Kerim Y. Aydin, and Robert C. Francis. 2011. What drives dynamics in the Gulf of Alaska? Integrating hypotheses of species, fishing, and climate relationships using ecosystem modeling. *Can. J. Fish. Aquat. Sci.* 68: 1553–1578.
- Guttormsen, M. A., and D. Jones. 2010. Results of the acoustic-trawl surveys of walleye pollock (*Theragra chalcogramma*) in the Gulf of Alaska, February-March 2010 (DY2010-01 and DY2010-02). AFSC Processed Rep. 2010-05, 85 p.
- Guttormsen, M. A., and P. T. Yassenak. 2007. Results of the 2003 and 2005 echo integration-trawl surveys in the Gulf of Alaska during summer, Cruises MF2003-09 and OD2005-01. AFSC Processed Rep. 2007-04, 61 p.
- Szalay, von, P.G.; N.W. Raring, F.R. Shaw, M.E. Wilkins, and M.H. Martin 2010. Data Report: 2009 Gulf of Alaska Bottom Trawl Survey. NOAA Technical Memorandum NMFS-AFSC-208. 257 p.
- Williams, K. Draft. Effect of trawl selectivity on the acoustic survey time series and assessment of Gulf of Alaska walleye pollock (*Theragra chalcogramma*). 40 p.
- Williams, K., Punt, A. E., Wilson, C. D., and Horne, J. K. 2011. Length-selective retention of walleye pollock, *Theragra chalcogramma*, by midwater trawls. *ICES Journal of Marine Science*, 68: 119–129.

## Appendix 2: Statement of Work

### External Independent Peer Review by the Center for Independent Experts

#### Gulf of Alaska (GOA) walleye pollock stock Assessment Review

**Scope of Work and CIE Process:** The National Marine Fisheries Service's (NMFS) Office of Science and Technology coordinates and manages a contract providing external expertise through the Center for Independent Experts (CIE) to conduct independent peer reviews of NMFS scientific projects. The Statement of Work (SoW) described herein was established by the NMFS Project Contact and Contracting Officer's Technical Representative (COTR), and reviewed by CIE for compliance with their policy for providing independent expertise that can provide impartial and independent peer review without conflicts of interest. CIE reviewers are selected by the CIE Steering Committee and CIE Coordination Team to conduct the independent peer review of NMFS science in compliance the predetermined Terms of Reference (ToRs) of the peer review. Each CIE reviewer is contracted to deliver an independent peer review report to be approved by the CIE Steering Committee and the report is to be formatted with content requirements as specified in **Annex 1**. This SoW describes the work tasks and deliverables of the CIE reviewer for conducting an independent peer review of the following NMFS project. Further information on the CIE process can be obtained from [www.ciereviews.org](http://www.ciereviews.org).

**Project Description** The Alaska Fisheries Science Center (AFSC) requests a Center of Independent Experts (CIE) review of the stock assessment for Gulf of Alaska (GOA) walleye pollock. The walleye pollock stock in the Gulf of Alaska is important to local fishing communities and is a key component of the GOA ecosystem. Walleye pollock stock assessments routinely undergo review by the AFSC, the North Pacific Fisheries Management Council's Groundfish Plan Team and Scientific and Statistical Committee. The assessment model for pollock has been stable for some time, and several significant changes are being contemplated for the 2012 assessment. In addition, the pollock stock assessment has not had the benefit of a CIE review since 2003. Therefore, a CIE review in 2012 would be timely. The Terms of Reference (ToRs) of the peer review are attached in **Annex 2**. The tentative agenda of the panel review meeting is attached in **Annex 3**.

**Requirements for CIE Reviewers:** Three CIE reviewers shall conduct an impartial and independent peer review in accordance with the SoW and ToRs herein. CIE reviewers must be thoroughly familiar with various subject areas involved in stock assessment, including population dynamics, survey methodology, and estimation of parameters in complex nonlinear models. Reviewers must also have experience conducting stock assessments for fisheries management. Expertise would be desirable in several other areas. First, since the pollock assessment uses AD Model Builder (ADMB) software, expertise in using this software would be desirable. Second, changes being considered for the 2012 assessment include adding ecological interactions and environmental forcing to the assessment model, so expertise in these areas would also be desirable. It is not expected that all three of the reviewers have these specialized areas of expertise, rather that at least one of the three reviewers should be knowledgeable in these areas. Each CIE reviewer's duties shall not exceed a maximum of 14 days to complete all work tasks of the peer review described herein.

**Location of Peer Review:** Each CIE reviewer shall conduct an independent peer review during the panel review meeting scheduled in Seattle, Washington with dates July 17-20, 2012.

**Statement of Tasks:** Each CIE reviewers shall complete the following tasks in accordance with the SoW and Schedule of Milestones and Deliverables herein.

Prior to the Peer Review: Upon completion of the CIE reviewer selection by the CIE Steering Committee, the CIE shall provide the CIE reviewer information (full name, title, affiliation, country, address, email) to the COTR, who forwards this information to the NMFS Project Contact no later the date specified in the Schedule of Milestones and Deliverables. The CIE is responsible for providing the SoW and ToRs to the CIE reviewers. The NMFS Project Contact is responsible for providing the CIE reviewers with the background documents, reports, foreign national security clearance, and other information concerning pertinent meeting arrangements. The NMFS Project Contact is also responsible for providing the Chair a copy of the SoW in advance of the panel review meeting. Any changes to the SoW or ToRs must be made through the COTR prior to the commencement of the peer review.

Foreign National Security Clearance: When CIE reviewers participate during a panel review meeting at a government facility, the NMFS Project Contact is responsible for obtaining the Foreign National Security Clearance approval for CIE reviewers who are non-US citizens. For this reason, the CIE 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 NMFS Project Contact for the purpose of their security clearance, and this information shall be submitted at least 30 days before the peer review in accordance with the NOAA Deemed Export Technology Control Program NAO 207-12 regulations available at the Deemed Exports NAO website: <http://deemedexports.noaa.gov/>  
[http://deemedexports.noaa.gov/compliance\\_access\\_control\\_procedures/noaa-foreign-national-registration-system.html](http://deemedexports.noaa.gov/compliance_access_control_procedures/noaa-foreign-national-registration-system.html)

Pre-review Background Documents: Two weeks before the peer review, the NMFS 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 NMFS Project Contact will consult with the CIE Lead Coordinator on where to send documents. CIE reviewers are responsible only for the pre-review documents that are delivered to the reviewer in accordance to the SoW scheduled deadlines specified herein. The CIE reviewers shall read all documents in preparation for the peer review.

Panel Review Meeting: Each CIE reviewer shall conduct the independent peer review in accordance with the SoW and ToRs, and shall not serve in any other role unless specified herein. **Modifications to the SoW and ToRs can not be made during the peer review, and any SoW or ToRs modifications prior to the peer review shall be approved by the COTR and CIE Lead Coordinator.** Each CIE reviewer shall actively participate in a professional and respectful manner as a member of the meeting review panel, and their peer review tasks shall be focused on the ToRs as specified herein. The NMFS Project Contact is responsible for any facility arrangements (e.g., conference room for panel review meetings or teleconference arrangements). The NMFS Project Contact is responsible for ensuring that the Chair understands the contractual role of the CIE reviewers as specified herein. The CIE Lead Coordinator can contact the Project Contact to confirm any peer review arrangements, including the meeting facility arrangements.

Contract Deliverables - Independent CIE Peer Review Reports: Each CIE reviewer shall complete an independent peer review report in accordance with the SoW. Each CIE reviewer shall complete the independent peer review according to required format and content as described

in Annex 1. Each CIE reviewer shall complete the independent peer review addressing each ToR as described in Annex 2.

**Specific Tasks for CIE Reviewers:** The following chronological list of tasks shall be completed by each CIE reviewer in a timely manner as specified in the **Schedule of Milestones and Deliverables**.

- 1) Conduct necessary pre-review preparations, including the review of background material and reports provided by the NMFS Project Contact in advance of the peer review.
- 2) Participate during the panel review meeting at the Seattle, Washington during July 2012 (dates to be determined by Project Contact no later than 15 April 2012).
- 3) In Seattle, Washington during 17-20 July 2012 as specified herein, and conduct an independent peer review in accordance with the ToRs (**Annex 2**).
- 4) No later than **August 3, 2012**, each CIE reviewer shall submit an independent peer review report addressed to the “Center for Independent Experts,” and sent to Mr. Manoj Shivlani, CIE Lead Coordinator, via email to [shivlanim@bellsouth.net](mailto:shivlanim@bellsouth.net), and CIE Regional Coordinator, via email to Dr. David Die [ddie@rsmas.miami.edu](mailto:ddie@rsmas.miami.edu). Each CIE report shall be written using the format and content requirements specified in Annex 1, and address each ToR in **Annex 2**.

**Schedule of Milestones and Deliverables:** CIE shall complete the tasks and deliverables described in this SoW in accordance with the following schedule.

18 June 2012	CIE sends reviewer contact information to the COTR, who then sends this to the NMFS Project Contact
3 July 2012	NMFS Project Contact sends the CIE Reviewers the pre-review documents
<b>17-20 July 2012</b>	Each reviewer participates and conducts an independent peer review during the panel review meeting
3 August 2012	CIE reviewers submit draft CIE independent peer review reports to the CIE Lead Coordinator and CIE Regional Coordinator
17 August 2012	CIE submits CIE independent peer review reports to the COTR
24 August 2012	The COTR distributes the final CIE reports to the NMFS Project Contact and regional Center Director

**Modifications to the Statement of Work:** This ‘Time and Materials’ task order may require an update or modification due to possible changes to the terms of reference or schedule of milestones resulting from the fishery management decision process of the NOAA Leadership, Fishery Management Council, and Council’s SSC advisory committee. A request to modify this SoW must be approved by the Contracting Officer at least 15 working days prior to making any permanent changes. The Contracting Officer will notify the COTR within 10 working days after receipt of all required information of the decision on changes. The COTR can approve changes to the milestone dates, list of pre-review documents, and ToRs within the SoW as long as the role

and ability of the CIE reviewers to complete the deliverable in accordance with the SoW is not adversely impacted. The SoW and ToRs shall not be changed once the peer review has begun.

**Acceptance of Deliverables:** Upon review and acceptance of the CIE independent peer review reports by the CIE Lead Coordinator, Regional Coordinator, and Steering Committee, these reports shall be sent to the COTR for final approval as contract deliverables based on compliance with the SoW and ToRs. As specified in the Schedule of Milestones and Deliverables, the CIE shall send via e-mail the contract deliverables (CIE independent peer review reports) to the COTR (William Michaels, via [William.Michaels@noaa.gov](mailto:William.Michaels@noaa.gov)).

**Applicable Performance Standards:** The contract is successfully completed when the COTR provides final approval of the contract deliverables. The acceptance of the contract deliverables shall be based on three performance standards:

- (1) The CIE report shall be completed with the format and content in accordance with **Annex 1**,
- (2) The CIE report shall address each ToR as specified in **Annex 2**,
- (3) The CIE reports shall be delivered in a timely manner as specified in the schedule of milestones and deliverables.

**Distribution of Approved Deliverables:** Upon acceptance by the COTR, the CIE Lead Coordinator shall send via e-mail the final CIE reports in \*.PDF format to the COTR. The COTR will distribute the CIE reports to the NMFS Project Contact and Center Director.

**Support Personnel:**

William Michaels, Program Manager, COTR  
NMFS Office of Science and Technology  
1315 East West Hwy, SSMC3, F/ST4, Silver Spring, MD 20910  
[William.Michaels@noaa.gov](mailto:William.Michaels@noaa.gov) Phone: 301-427-8155

Manoj Shivlani, CIE Lead Coordinator  
Northern Taiga Ventures, Inc.  
10600 SW 131<sup>st</sup> Court, Miami, FL 33186  
[shivlanim@bellsouth.net](mailto:shivlanim@bellsouth.net) Phone: 305-383-4229

Roger W. Peretti, Executive Vice President  
Northern Taiga Ventures, Inc. (NTVI)  
22375 Broderick Drive, Suite 215, Sterling, VA 20166  
[RPerretti@ntvifederal.com](mailto:RPerretti@ntvifederal.com) Phone: 571-223-7717

**Key Personnel:**

NMFS Project Contact:

Martin Dorn et al.  
NOAA National Marine Fisheries Service  
Alaska Fisheries Science Center  
7600 Sand Point Way, NE, Seattle, WA 98115-6349  
Email: [martin.Dorn.et.al.@noaa.gov](mailto:martin.Dorn.et.al.@noaa.gov) Phone: 206-526-6548

## **Annex 1: Format and Contents of CIE Independent Peer Review Report**

1. The CIE independent report shall be prefaced with an Executive Summary providing a concise summary of the findings and recommendations.
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. Reviewers 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. Reviewers 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. Reviewers shall provide a critique of the NMFS review process, including suggestions for improvements of both process and products.
  - d. The CIE independent report shall be a stand-alone document for others to understand the weaknesses and strengths of the science reviewed.
3. The reviewer report shall include the following appendices:
  - Appendix 1: Bibliography of materials provided for review
  - Appendix 2: A copy of the CIE Statement of Work
  - Appendix 3: Panel Membership or other pertinent information from the panel review meeting.

**Annex 2: Terms of Reference for Peer Review of the  
Gulf of Alaska Walleye Pollock Stock Assessment**

1. Evaluate and provide recommendations on data collection procedures and analytical methods used to develop assessment model input.
2. Evaluate and provide recommendations on model structure, assumptions, and estimation procedures.
3. Evaluate and provide recommendations for the reporting of assessment results and characterization of uncertainty.
4. Evaluate and provide recommendations on F35% spawning biomass per recruit as an appropriate proxy for FMSY under non-stationarity in vital rates. Also evaluate and provide recommendations on the B35% biomass reference point as a proxy for BMSY.
5. Recommendations for further improvements.
6. Brief description on panel review proceedings highlighting pertinent discussions, issues, effectiveness, and recommendations.

### **Annex 3: Tentative Agenda**

#### **Review of the Gulf of Alaska Walleye Pollock Stock Assessment**

Alaska Fisheries Science Center, NOAA  
7600 Sand Point Way N.E., Building 4  
Seattle, Washington 98115  
Phone: 206 526-4000  
17-20 July 2012

The final meeting agenda has not yet been drafted, but will be forwarded by the project contact as soon as it becomes available.

17 July 2012	Presentations by survey and fishery data collection scientists
18 July 2012	Presentation by assessment scientists, Panel discussion and requests
19 July 2012	Panel discussion and requests, Begin drafting reviewer reports
20 July 2012	Draft reviewer reports

### **Appendix 3: Panel membership and meeting agenda**

The review panel consisted of three CIE appointed reviewers:

Mr. Patrick Cordue, Fishery Consultant, New Zealand  
Dr. Carmen Fernández, Vice-Chair, ICES Advisory Committee, Spain.  
Dr. Ian Jonsen, Dalhousie University, Canada

The meeting was chaired by Dr. Hollowed. The draft agenda (below) was closely followed on Tuesday, Wednesday, and Thursday. On Friday, exploration of a working model was continued.

#### **Review Panel Meeting on Gulf of Alaska Stock Assessment Draft Agenda**

July 17-20, 2012  
Alaska Fisheries Science Center  
7600 Sand Point Way NE, Seattle, WA 98112

---

#### **Tuesday, July 17, 2012**

9:00 a.m.	Welcome and Introductions, Adopt Agenda	Anne Hollowed
9:15 a.m.	Overview of biology, surveys, fishery, management system	Martin Dorn et al.
10:00 p.m.	Gulf of Alaska bottom trawl survey	Michael Martin 1 hr
11:00 p.m.	Acoustic surveys in the Gulf of Alaska	Mike Guttormsen/Chris Wilson 1 hr
12:00 p.m.	Lunch	
1:30 p.m.	Evaluation of net selectivity in acoustic surveys	Kresimir Williams 1 hr
2:30 p.m.	Fishery monitoring of the GOA pollock fishery	Martin Loefflad or alternate 1 hr
3:30 p.m.	Role of pollock in the GOA ecosystem	Kerim Aydin 1 hr
5:00 p.m.	Meeting adjourns for the day	

#### **Wednesday, July 18, 2012**

9:00 a.m.	Morning welcome and announcements	
9:15 a.m.	Pollock stock assessment model	Martin Dorn et al. 3 hrs
12:00 p.m.	Lunch	
1:30 p.m.	Management Strategy Evaluation of GOA pollock assessment	Teresa A'mar 2 hr
3:30 p.m.	Discussion of proposed assessment model changes	Martin Dorn et al. 2 hr
5:00 p.m.	Meeting adjourns for the day	

**Thursday, July 19, 2012**

9:00 a.m. Morning welcome and announcements

9:15 a.m. Evaluation of alternative model configurations

12:00 p.m. Lunch

1:30 a.m. Continued evaluation of alternative model configurations

**Friday, July 20, 2012**

9:00 a.m. Report writing. AFSC analysts will be available to respond to requests and to answer questions

## Appendix 4: Acoustic proportion-at-age examples

These examples are for illustrative purposes only. They are not an attempt to estimate the biases inherent in the proportion-at-age estimates from the acoustic surveys. They simply demonstrate circumstances under which biases would occur.

### Example 1: Size/age based vulnerability to trawl gear

Suppose that there is a spatially homogeneous layer of mixed-age-class pollock which after post-stratification is placed in a single length-stratum. Suppose that the layer is referenced to the bottom and extends 100m above the seabed. Further, assume that there is vertical stratification by size (smaller fish in mid water, larger fish nearer the bottom) and that the fish have a length/age specific dive/avoidance reaction. So, in addition to a mesh selection (smaller fish are more likely to go through the meshes than larger fish) there is also an avoidance selection (larger fish avoid the net more than smaller fish) and a vertical-availability selection (smaller fish are higher in the water column than larger fish). Clearly, the expected proportion-at-age in the net will differ from the proportion-at-age in the layer.

As an illustration, suppose that trawling was done using mid-water gear near the top of the layer. An assumed selection pattern and true proportion-at-age is given in the table below together with the expected proportion-at-age in the net.

	<b>Age 1</b>	<b>Age 2</b>	<b>Age 3</b>	<b>Age 4+</b>
Relative selectivity	0.2	0.8	0.9	0.1
Proportion-at-age in layer	0.50	0.20	0.10	0.20
Expected proportion-at-age in the trawl	0.27	0.43	0.24	0.05

If the trawling was done with bottom gear, a different selection pattern would apply and a different expected proportion-at-age in the net would occur as illustrated in the table below.

	<b>Age 1</b>	<b>Age 2</b>	<b>Age 3</b>	<b>Age 4+</b>
Relative selectivity	0.1	0.6	0.9	0.8
Proportion-at-age in layer	0.50	0.20	0.10	0.20
Expected proportion-at-age in the trawl	0.12	0.29	0.21	0.38

If the true mesh-selectivity was known for the 1-year old fish and a correction was applied it would move the expected proportion-at-age closer to the proportions in the layer in these examples (because there is a high proportion of 1 year-old fish):

	<b>Age 1</b>	<b>Age 2</b>	<b>Age 3</b>	<b>Age 4+</b>
Proportion-at-age in layer	0.50	0.20	0.10	0.20
Corrected expected proportion-at-age in the midwater trawl	0.65	0.21	0.12	0.03
Corrected expected proportion-at-age in the bottom trawl	0.57	0.14	0.10	0.18

### **Example 2: Spatial structure**

Suppose there are two homogeneous mixed-age-class pollock layers which cover adjoining regions. They look like a continuous layer but have different age-class mixes and contain different numbers of fish:

	<b>Age 1</b>	<b>Age 2</b>	<b>Age 3</b>	<b>Age 4+</b>
Layer 1 (80 million fish)	0.50	0.20	0.10	0.20
Layer 2 (20 million fish)	0.20	0.40	0.40	0.00
Combined layer proportion-at-age	0.44	0.24	0.16	0.16

Also, layer 1 has moderately dense marks and covers a large area, while layer 2 has denser marks over a smaller area.

Suppose, in the first instance, that (yet to be invented) sampling gear with a uniform selection pattern is used (i.e., all age classes are equally selected). If we know the probabilities of samples being taken from the different layers and the probability of post-stratification into separate length-strata, then we can calculate the expected proportion-at-age for the sampling gear and design.

There are four probabilities that we need to assume for the calculation:

<b>Scenario</b>	<b>Probability</b>
Only layer 1 sampled (single length-stratum)	0.1
Only layer 2 sampled (single length-stratum)	0.6
Samples from both layers; single length-stratum	0.1
Samples from both layers; two length-strata	0.2

Sampling from layer 2 only is given the highest probability because it has the densest marks.

We can also calculate the expected proportion-at-age for gear with any given selection pattern. The table below compares the true proportion-at-age (in the combined layer) with the expected proportion-at-age when there is no selection pattern and when the mid-water gear of Example 1 is assumed.

	<b>Age 1</b>	<b>Age 2</b>	<b>Age 3</b>	<b>Age 4+</b>
Combined layer proportion-at-age	0.44	0.24	0.16	0.16
No selection pattern	0.29	0.34	0.31	0.06
Mid-water gear	0.10	0.44	0.45	0.01

## Appendix 5: Model equations and methods used for the evaluation of $F_{35\%}$ and $B_{35\%}$

A simple, numbers-only model, with a modified Baranov catch equation was used to evaluate the suitability of  $F_{35\%}$  for a population where the dynamics are driven by predator abundance. The model is theoretical and is not meant to be a representation of the GOA Pollock stock. It is intended to capture important differences in the dynamics of a single-species stock with constant natural mortality compared to one with variable natural mortality driven by predator abundance.

The model keeps track of the numbers of mature fish in a single-species population. No age structure or growth is modelled. There is an annual cycle which consists of recruitment then a full year of mortality from three sources: a constant “base” natural mortality ( $M$ ), additional natural mortality from specified predation rates ( $P$ ), and a constant fishing mortality ( $F$ ). A Beverton-Holt stock-recruitment relationship is assumed with a lag of four years (e.g., fish spawned in 1980 recruit to the mature population in 1984).

The catch equation is a piecewise function, being the standard Baranov catch equation when numbers in the population are below  $N_c$  (the “c” stands for “constrained”) and a modified form when the numbers are above  $N_c$ . The modified equation derives from the following differential equation which has unconstrained predation:

$$\frac{dN}{dt} = -(M + F)N - P$$

where the population number  $N$  is a function of time  $t$  and  $P$  is the total predation rate due to a number of predator species:

$$P = \sum_i P_i$$

For the purposes of solving the equation,  $P$  is a constant (the total predation rate in the model is specified each year). The solution to the differential equation is:

$$N(t) = N_0 e^{-Zt} - \frac{P}{Z} (1 - e^{-Zt})$$

where  $N_0 = N(0)$  and  $Z = M + F$ .

The removals, from time zero through to time  $t$ , due to the specified predation rate is simply  $Pt$ . For fishing and natural mortality the removals are respectively:

$$\frac{F}{Z} \left[ \left( N_0 + \frac{P}{Z} \right) (1 - e^{-Zt}) - Pt \right]$$

and

$$\frac{M}{Z} \left[ \left( N_0 + \frac{P}{Z} \right) (1 - e^{-Zt}) - Pt \right]$$

These equations assume that the numbers are sufficient in the (prey) population to allow the predators to eat as much as they need – i.e., consumption is unconstrained by prey abundance. When population numbers are below  $N_c$  it is assumed that consumption is constrained by prey numbers:

$$\frac{dN}{dt} = -(M + F)N - kPN$$

where  $k$  is a constant to be determined so that there is a “seamless” transition from one catch equation to the next. That is, the derivatives must be equal when  $N(t) = N_c$ :

$$-(M + F)N_c - P = -(M + F)N_c - kPN_c$$

and hence,  $k = \frac{1}{N_c}$ .

Therefore, when population numbers are below  $N_c$  the catch equation is just Baranov with three constant instantaneous mortality rates,  $M$ ,  $F$ , and  $\frac{P}{N_c}$ .

In a year (within the model) when there is a transition from unconstrained to constrained predation it occurs at time  $t$ :

$$t = \frac{1}{Z} \ln \left[ \frac{N_0 + \frac{P}{Z}}{N_c + \frac{P}{Z}} \right]$$

Since the cycle is annual, this means that in years when there is a transition, the unconstrained equations apply for duration  $t$  and the constrained equations apply for duration  $1 - t$ .

It was assumed that in the virgin state the population was in “stochastic” equilibrium. The initial numbers were set equal to 100 and virgin recruitment was determined so that the average annual (beginning of year) population number over a run of 1000 years was equal to 100.

Predator removal rates over the 1000 year duration were specified as sine functions with different amplitudes ( $a_i$ ), periods ( $p_i$ ), and average removal rates ( $P_{i,av}$ ):

$$P_i = P_{i,av} \left[ a_i \sin \left( \frac{2\pi}{p_i} t \right) + 1 \right] \quad t = 1, 2, \dots, 1000$$

In the runs which used specified predation, three predators with the following parameters were used:

Predator	Average rate ( $P_{i,av}$ )	Amplitude ( $a_i$ )	Period ( $p_i$ )
1	10	0.5	23
2	2	-0.5	17
3	5	0.8	31

The prime number periods were used so that the total predation rate would have a very long period – so that the predation pattern does not (exactly) repeat during the 1000 years of each run.

$F_{MSY}$  and related reference points and statistics were calculated for six model runs consisting of three levels of steepness (0.7, 0.8, 0.9) combined with two values of  $N_c$  (10, 50). Yields and other statistics (over a range of  $F$  values to generate a yield curve) were calculated from the average of annual values from years 51 to 1000 (i.e., a burn-in of 50 years was allowed for the model to reach “equilibrium”).

The value of  $F_{35\%}$  was determined for an “equivalent” constant- $M$  model without specified predation. That is, predation was set to zero (while virgin recruitment was unchanged) and natural mortality was increased to the level that maintained population numbers at virgin levels when there was no fishing. The value of  $B_{35\%}$  (or  $N_{35\%}$  since the models are numbers only) was calculated for the constant- $M$  model and the six predation models (using runs of 1000 years and average statistics as above).