Review of the Gulf of Alaska walleye pollock (*Theragra chalcogramma*) assessment.

Malcolm Haddon

33 Jindabyne Road
Kingston
Tasmania 7050
Australia

Marine Research Laboratory
Tasmanian Aquaculture and Fisheries Institute
University of Tasmania
Nubeena Crescent, Taroona
Tasmania 7053
Australia

Malcolm.Haddon@utas.edu.au

Executive Summary

The Gulf of Alaska pollock surveys, assessment and management appear to be well structured, are relatively well developed and have only a few unexplored options. Because of the variable ecosystem of which it is a part, the pollock fishery is an exceptionally difficult species for which to generate an assessment with a high degree of certainty. Many of the uncertainties in the surveys and the assessment lead to a conservative assessment (e.g. the survey estimates are often biased low, the constant natural mortality for all ages biases the spawning biomass low), which leads to the conclusion that the current management is precautionary. The fishery is, in fact, only a small part of the dynamics of the stock. It is possible the management triggers are overly rigid and do not recognize that the assessment is made in the presence of significant uncertainty. Thus, while the assessment may eventually suggest that the limit reference point of 20% unfished spawning biomass has been passed, it can only claim this with a given degree of certainty. The current management is certainly precautionary in the context of the present assessment to stock trends and its associated uncertainties. It is recommended that rather than using an abrupt decision rule based on an uncertain assessment, some form of risk assessment be developed instead that suggests management responses in the light of the varying likelihood of different events occurring. Some way of accounting for the uncertainty inherent in the assessment should be included in the management decision process.

As discussed in this document there are actions that could be adopted which may reduce the uncertainty in the stock assessment of the Gulf of Alaska pollock. Unfortunately, there are strong indications that the ecosystem in the Gulf of Alaska is unstable and a number of significant species, including pollock, are undergoing long-term and large trends in population size. This is perhaps the greatest source of uncertainty in the pollock assessment, but, other than monitoring the other species (especially arrowtoothed flounder and Steller sea-lions) and investigating the interactions between the species, there is little that the current assessment can do to integrate such factors. Multi-species models could be investigated but any that assume equilibrium, even as a starting point, are unlikely to provide any useful insights.

We have made a number of suggestions that might lead to improved estimates of biomass or at least of the uncertainty associated with the assessment. These are listed through the document and a summary listing is given in the conclusions. The easiest to implement and the highest priority issues include:

The time-series of bottom trawl abundance estimates is really a combination of two time-series with either a step change between them or something more gradual. It is suggested that either this is demonstrated to have only a minor effect or this be recognized in the modelling by adding the required parameters to characterize the different time series as a better approximation to reality than the current design.

A longer term aim should be to move the design of the bottom trawl surveys to a fixed station design.

The current swept area estimates include untrawlable ground, which assumes that pollock occur at the same average density over the foul ground as it does over the trawlable ground. It would be relatively simple to compare the outcomes of the surveys with and without this untrawlable area.

The strategy of expanding the geographical coverage of the Echo-integration trawl (EIT) surveys should be continued with surveys of other spawning grounds being included. The three valid Shumagin EIT surveys should be included in the current assessment.

Some way should be developed for including estimates of biomass from the dead-zone close to the seabed in the EIT surveys.

A high priority should be given to the effects of providing length and weight ancillary information to the people ageing otoliths. Removal of this interdependence between age and length would improve the model estimation process by keeping the different likelihoods strictly independent.

The ageing error matrix should be updated using the latest information.

A high priority should be given to the histological investigation of gonad samples to clarify the current remarkable variability in the observed size-at-maturity.

The stock assessment model should be modified to investigate the influence of adding the one-year old fish to the dynamics along with alternative schedules of natural mortality by age. This, combined with the investigation of the relative consistency of the various data sources, by their systematic exclusion, along with a determination of their relative contribution to the overall uncertainty (as perhaps expressed by the width of the likelihood profiles) would enable a more defensible assessment to be written.

Background

Statement and History of the Problem

Walleye pollock (*Theragra chalcogramma*) in the Gulf of Alaska have declined from a maximum abundance in the late 1970s/early 1980s to a 30 year low in 2002, despite repeated reductions in the Total Allowable Catch (TAC) being made. Pollock abundance estimates are produced from both acoustic and trawl surveys, and the results of these are used as inputs to a catch-at-age model that also uses ancillary information (termed a Stock Synthesis or Integrated Analysis model). Confidence in both the results of some recent survey results, and the assessment, has been reduced because of four influences: 1) the echo integration-trawl and the bottom trawl surveys producing somewhat contradictory trends, 2) a marked increase in the uncertainty around the more recent bottom trawl surveys, 3) evidence for apparent changes in the spawning stock distribution (especially affecting the echo integration-trawl survey), and 4) changes in the observed values of some important biological characters of the Gulf of Alaska pollock (size at maturity, length and weight at age). A reliable assessment of Gulf of Alaska pollock stock status is important both to allow the North Pacific Fishery Management Council to make informed decisions about this fishery and for evaluating fishery impacts on stocks of the endangered Steller Sea Lion, for which pollock are important prey.

The Gulf of Alaska walleye pollock fishery is, in many ways, biologically distinct from the Bering Sea pollock fishery. Genetically, the evidence is mixed concerning the relationship between the two areas; some sources find no significant differences, while others find minor but significant differences. Whatever the case may be, the dynamics of the Gulf of Alaska pollock appear to be sufficiently disarticulated from the populations in the Bering Sea that separate management is essential. Large changes have developed in the Gulf of Alaska coastal shelf system, including the long-term decline of Steller Sea Lions, the long-term decline of the Walleye pollock, and the long-term increase in Arrowtooth flounder. These species are mentioned explicitly because Gulf of Alaska pollock are reportedly a major prey item for both the presently endangered Steller Sea Lion and the Arrowtooth Flounder. The possibility that the Steller Sea Lion stocks are suffering as a result of the decline in the pollock has raised serious concerns and has focussed attention on the Gulf of Alaska pollock assessment and its consequent management.

The assessment is based around a number of sources of data including summer bottom trawl surveys conducted by the NMFS, summer bottom trawl surveys conducted by the ADFG, winter echo integration-trawl surveys, the fishery, earlier egg-production surveys and various ancillary biologically based data sources. The management strategy is set as a function of the spawning biomass so details such as the age-structure and size-at-maturity for the pollock also constitute an important foundation to the assessment. This review will begin with a consideration of all of the data inputs with their strengths and weaknesses, this will be followed by a consideration of other data that could be used in the assessment, and finally by a consideration of the assessment model and its potential for providing management advice. The review is based upon two sources of information. First, an extensive literature provided by the Alaska Fisheries Science Center (AFSC) both before and during the review process (see References section) and second, the review process itself, based at the AFSC in Seattle, which consisted of presentations by staff followed by detailed discussions.

Review Activities

The review was conducted at the Alaska Fisheries Science Center in Seattle over the five days: Monday 4th August to Friday 8th August 2003. The review process followed a format of brief presentations by the various contributors to the Gulf of Alaska pollock assessment with consequent questions and discussions. The timetable of presentations and discussions was:

Monday 4th August
Dr Martin Dorn - Overview
Dr Bernard Megrey - Reproductive Biology/Size at Maturity
Dr Chris Wilson - Echo Integration Trawl Surveys

Tuesday 5th August
Mark Wilkins - NMFS Bottom Trawl Surveys
Paul von Szalay - Alaska Department of Fish and Game Trawl Surveys

Wednesday 6th August
Dr Kevin Bailey - Stock Structure
Dr Dan Kimura - Ageing of pollock
Pat Livingston, Sarah Gaichas - Ecological Interactions
Dr Anne Hollowed, Dr Jim Ianelli - Harvest Policy Management

Thursday 7th August
Dr Bill Karp, Martin Loefflad - Fishery Monitoring and Sampling
Dr Martin Dorn - Assessment Model

Friday 8th August
Informal meetings as necessary.

The emphasis of the review was on:

1. Whether the current management approach is precautionary, given current stock status and trends;

2. Strengths and weaknesses of the current surveys for assessing Gulf of Alaska pollock abundance trends (adequacy of existing surveys to monitor abundance trends of GOA pollock);

3. The strengths and weaknesses of the Gulf of Alaska pollock assessment and harvest strategy (whether the stock assessment takes into account major uncertainties in data and assumptions);

4. Recommendations for improvement to the assessment model and harvest strategy;

5. Recommendations for survey improvements to better assess Gulf of Alaska pollock;

6. Suggested research priorities to improve the stock assessment.

In line with the review terms of reference, Dr Godo was to focus more attention on the survey designs and adequacy, while I was to pay extra attention to the assessment and modelling. Meeting with the scientists involved in the research at the AFSC in Seattle was necessary because the range of inputs to the assessment model, and the fact that this is an on-going evolving process, meant that simple documentary material would be insufficient to provide a sufficient grasp of the issues and their present status. It eventuated that some suggestions made by the reviewers are already in the process of being adopted or investigated for adoption within the assessment process.

The review process was conducted in a positive and friendly atmosphere with great interest and enthusiasm being expressed by the assessment team in Seattle for any discussion relating to their work. The Gulf of Alaska pollock assessment is a complex and difficult problem so it is fortunate that the researchers in Seattle are so enthusiastic and committed. Their openness to critical discussion does them and their organization credit.

I would like to thank the assessment team at the AFSC for making the review such an interesting and positive experience.

DISCLAIMER

The information in this review has been provided by way of review only. The author makes no representation, express or implied, as to the accuracy of the information and accepts no liability whatsoever for either its use or any reliance placed on it.

Summary of Findings

Structure of Document

This review will be structured to parallel, approximately, the data inputs to the assessment model. Thus, there will be a discussion of each of the survey types: NMFS trawl survey, ADFG trawl survey, NMFS echo integration-trawl (EIT) survey. This will be followed by a general discussion concerning biological inputs to the assessment model, including the numbers-at-age data, the size-at-maturity data, other biological data, and finally the fishery data. There is also a discussion concerning the model structure and implementation with commentary on how well it informs the management decisions.

NMFS Bottom Trawl Surveys

These surveys are conducted, for operational reasons, in the summer months and are multi-species surveys optimised as a compromise between many species using standard design methods. For such objectives, these surveys are well designed. While there have been a multiplicity of vessels used, the Alaskan Fisheries Science Center have been attempting to initiate multi-year charters to ensure the use of the same vessels and gear. If formal comparison trials cannot be made between vessels then it is suggested that the fishery data from the vessels used be standardized for at least area and month, using GLMs, in an attempt to determine whether there are marked differences between the vessels. Ideally, vessel comparison trials should be conducted, but fishery standardization procedures may provide some insights provided the vessels have fished in roughly the same areas over the same periods, using approximately the same gear.

Currently the surveys use three vessels for about 2.5 months during May to August. The depth stratifications match the distribution of pollock reasonably well. The strata are also determined by the distinct habitats present in the survey area, which is a sensible strategy in a multi-species survey over such diverse sea bed structure.

The swept area calculations are well standardized by the use of custom designed and built gear that informs when the gear is actually on the bottom and fishing. This provides a very accurate bottom time and distance. In addition, it has also permitted the standardization of how the different skippers terminate a tow. Effort data is relatively clean since the 1996 survey when this gear was first used. The use of this custom built equipment is an excellent innovation.

The introduction of this bottom sensing gear was not the only significant change in the time-series of the Bottom Trawl Survey (BTS). In the surveys from 1984 to 1993, the tows were 30 minutes long; but from 1996 onwards, they were only 15 minutes long. This change was made primarily so that the surveys could sample areas of rougher bottom, which only permitted access for shorter tows. Without the improvement to knowing when the gear was on the bottom the shorter tows could have introduced significant variation in actual fishing time. The longer tows prior to 1996 would proportionally have helped reduce the impact of the lack of knowledge about actually bottom time. Because the tows were on the bottom longer than previously thought, the catch-per-unit-effort was inflated pre-1996. This has yet to be corrected in the assessment, thus current data, pre-1996, is biased high.

While the changes to the fishing procedures were entirely justified and understandable, they do imply that there is a significant break in the time series of the BTS used in the Gulf of Alaska pollock assessment. Whether it was a step change or more gradual was discussed in the review, but the important thing is that it has not been constant, as is currently assumed in the assessment model. This problem could be addressed by modifying the necessary parameters in the model either as a step function of a linear change for the same number of parameters.

The people conducting the surveys were very pleased with the acknowledgement of the limits and assumptions behind the surveys made by the modellers. Everyone appeared to be well versed with the limitations of each of the Bottom Trawl surveys. These were:

- Multi-species – optimised for numerous species, not just pollock.
- Availability of habitat – some habitats were unavailable because they were untrawlable.
- Patchy species vs. swept area calculations – inherently difficult to obtain low variance estimates of density.
- Trawl selectivity – this is not a constant with depth as the trawl mouth alters shape depending on the ratio of warps to depth used.
- Vertical availability – not all pollock are available on the bottom sampled by the bottom trawl.

It was recommended during the review that the surveyors consider using fixed stations, as this would improve their efficiency enormously (no need to search for suitable tow paths, nor for cruising over the expected tow path looking for foul ground) and permit many more stations to be visited. As long as the initial allocation were random the survey would still produce a workable relative abundance index. One response in the review was that the surveys also have the objective of characterizing the sea-bed of the Gulf in terms of foul ground and habitat types. It is considered that once a better perception of the Gulf sea-bed structure is obtained then they may move to a fixed survey site location design. This appeared to be a relatively inefficient way of characterizing the sea-bed in the Gulf. Perhaps it would be better to use the acoustic surveys across the Gulf to map the bottom types with the extent of different habitats. The benefits to be obtained from a fixed station survey, in terms of increasing the number of stations visited and the full recovery of data (no need to remove electronic gear because of marginal or risky tows), should encourage the move towards such a design as soon as possible. This, of course, depends upon the importance of the objective of determining the bottom types and whether an alternative method of mapping can be developed.

Currently, the large areas of foul ground are included in the calculations of biomass from swept area. This inclusion obviously assumes that the surveyed species live over the unfishable ground at densities equal to that over the fishable areas. For many bottom dwelling species this is a doubtful assumption and constitutes a significant source of uncertainty that is currently unaccounted for in the assessment. Whether it biases the survey estimates high or low in the case of pollock is unknown.

The present surveys use a set of standard warp : depth ratios, which implies that the selectivity is likely to change with depth. Fortunately, the survey area is stratified by depth so the surveys still constitute an index of relative abundance (instead of absolute abundance). Dr Godo suggested the use of a spreader limitation device as a means of standardizing the spread ratio of the trawl mouth. However, this would constitute another significant alteration in the time series; so perhaps, if it is to be adopted, it should wait and be included in a suite of other changes desired for the surveys (perhaps adopt this at the time of switching to a fixed site design).

NMFS Echo Integration-Trawl Surveys

The winter Echo Integration-Trawl (EIT) Surveys are primarily focussed upon the spawning stocks that occur within the Shelikoff Strait. During the review, there was discussion over whether an active turnover of stock within the Strait occurred. However, there were no reported changes to the average weights of females or to the sex-ratio during the brief two week spawning season, which suggests that if real turnover was occurring it was only trivial.

A major issue is that it is unknown exactly what proportion of the total stock is surveyed each winter by the EIT survey in the Shelikoff Strait. The spawning stock present is not necessarily a constant proportion of the total spawning stock or of the total stock. An estimate of this is made during the assessment modelling and it is the fact that this proportion appears to have altered markedly in recent years that has raised concerns about how well the EIT surveys represent the total stock.

The assessment team are well aware of this problem and are moving to adopt a number of positive approaches designed to reduce its influence. The first is to expand the geographical extent of the survey in the winter. There can be operational problems with such an expansion as witnessed by only three of the surveys of the Shumagin area being conducted at a time when the main spawning was occurring. It is recommended that these three surveys should be included in the assessment as they appear to have produced valid estimates of the spawning biomass present at the time of the surveys and they should complement the Shelikoff Strait estimates. In addition, the further expansion of the winter survey should be considered, if possible. The second innovation is to conduct EIT surveys in the summer to complement the summer NMFS bottom trawl surveys. This is undoubtedly a resource hungry approach but, because the survey is more general in its geographical coverage, there is a greater chance that it would provide a better estimate of the overall biomass (although there would be more problems of species identification than in the winter spawning survey).

Not finding all of the spawning aggregations during the present acoustic surveys obviously leads to a conservative estimate of spawning biomass. This implies that the earlier estimates, with their geographically more constrained scope would under-estimate productivity. The impact this would have on the predictions and hind-casting of the modelling are not simple, because the effect would not have been constant through time. However, the general effect would be to under-estimate total productivity and hence lead to overly conservative management.

The acoustic configuration has not been constant through time but, even though the vessel characteristics have been noisy, there have been no obvious signs of bias arising from these sources.

The major identified problem areas in the EIT surveys included a) species assignments, b) the length frequency weighting procedure, c) adjustment for other species, and d) the relationship between pollock Target strength to length of fish. Dr Godo will provide more detailed discussion of the survey issues, however, a few comments can be made here.

Eulachon (*Thaleichthys pacificus)*, a species that mixes with pollock in the winter, has no swim bladder so by using the same target strength as for pollock to account for the presence of eulachon, the EIT survey is producing a conservative estimate of pollock biomass. There is the possibility of post-processing this information to improve the estimates of pollock spawning biomass. Leaving the eulachon in the analysis as at present will not increase the uncertainty, but it will introduced a downward bias, albeit slight, into the assessment.

Apparently, species heterogeneity is not a large problem in the Shelikoff Strait survey but is likely to become more of a problem if a summer survey is adopted across the Gulf of Alaska. Presently, a bigger problem in the winter is a heterogeneity in size composition, but improvements with the estimates of the pollock target strength to length relationship should improve this situation. Dr Godo suggested that a lowered transducer, placed closer to the fish at depth would provide a better estimate of individual target strength.

Where else to look for spawning aggregations in the winter surveys could be suggested by the mapping of the older fishery data and the Joint Venture fishery data. During the review, new and detailed maps of the old joint venture fishery data were displayed by Dr Ianelli and these appear to provide good indications of older areas of spawning. They also suggested that some of the old data derived from outside the Gulf of Alaska. This latter circumstance should be investigated further and the fishery statistics adjusted appropriately if necessary.

A further source of conservative estimates would be how the dead-zone close to the sea bottom is handled. It was suggested that depth profiles of biomass would suggest how best to approach a dead-zone correction and make for a more defensible assessment.

Alaskan Department of Fish and Game Bottom Trawl Surveys

These surveys are primarily a survey of crab populations with detailed notice taken of the major bycatch species, which includes pollock. They operate with a different net and in a different habitat (they only survey smooth bottomed embayments, especially around Kodiak Island). Their net is relatively low opening; experiments put it as approximately 4x less efficient than the NMFS survey nets at catching pollock, which is probably related to the vertical opening height. All stations are fixed in the survey design, and the tows are run at a constant depth, for 1NM at 2knots. These surveys are usually run about 1 month later than the NMFS BT survey, though there is some overlap with the later parts of the NMFS survey.

Some details of the ADFG surveys need clarification in relation to how thorough the collection of data on the fish bycatch was in earlier years. If this was more *ad hoc* in earlier years, then less confidence should be placed in that data.

Once again it is uncertain whether the proportion of the stock represented in the embayments surveyed is a constant. Obviously if it were variable, then it would not be a good index of relative abundance. At the same time, if the embayments are the last to see a decline during the summer, that is they are a preferred habitat and would be the last to see a depletion, then it is also possible that they are hyper-stable and provide a biased view of stock status through time. If the ADFG data is to be included in the assessment, it seems necessary to determine whether the different data sets are giving essentially consistent information about any variations in stock status. It is suggested that a pseudo-jackknife procedure be implemented whereby each of the various surveys be removed from the analysis in turn and the relative impact determined. This would only constitute a small expansion of the sensitivity testing already being conducted by Dr Dorn, and should provide for a more defensible assessment.

General Comments on the Surveys

Dr Godo suggested the use of a forward-looking net-sonar to test the vertical distribution of pollock as perceived by the trawl. This is certainly worth considering for the EIT survey but it would also provide an estimate of the proportion of pollock available to the bottom trawl surveys; which is currently an unknown.

If there are to be configuration changes to the trawl surveys, it would be a reasonable idea to apply these to the EIT surveys first. In this way, it may be possible to generate a selectivity curve of each net and correct backwards through time.

Otherwise, the surveys and their design appear to have attained a degree of sophistication that means they should achieve their objectives as well as can be expected for line transect surveys, aimed at multiple species, at this geographical magnitude.

Numbers-at-Age Data

The assessment model is essentially a catch-at-age model with ancillary data, so the importance of the ageing data cannot be over-emphasized.

Fortunately, the criteria that have been used to age pollock have remained essentially unchanged since 1981. From the 1970s through to 1981, age estimates were made from whole otoliths. From 1981 to the present, the standard method is the classic break and burn approach, following a whole otolith examination. Two workshops have been held and both concluded that for pollock the break and burn method provides the best age estimates (Anon, 1990; Anon, 1998). Apparently, the ageing group at the AFSC are investigating the use of radio-isotopes to verify the ageing, especially of the older animals.

There was some discussion over sample sizes, the advantages and otherwise of random versus stratified sampling, and whether to sample fewer tows more thoroughly or sample as many tows as possible. This discussion was partly fuelled by the relatively low numbers of otoliths aged for the EIT surveys in recent years. Whatever the sampling scheme finally adopted, the minimum number of otoliths for an adequate sample should be defined; whether this should be 400 or 500 or even more will depend on the variation observed and the quality of the resulting age-length key.

The only area in the ageing routines that might be cause for concern was the provision of ancillary data concerning the length and weight of fish for each otolith aged. The potential for this extra information leading to biased answers is reasonably well established. At the very least it implies that the information about age is confounded, at least for the younger, smaller fish, with the length information; thus if modal classes exist in the lengths, these could well be considered to be specific ages, in this way the two data sources become confounded.

If the practice of provided the lengths and weights of fish to be aged is to continue, then it would be more defensible if a study was conducted that compared the age estimates obtained with and without this extra information. If there is a confounding between the length and age data, then the expectation would be that the precision of the younger age classes would change more than that for the older age classes (for which the length information would be less informative, though the weights might still provide information suggestive of age). In addition to the absolute age estimates, it is suggested that the error estimates with bias and between reader differences be characterized during the same tests.

At present there is only one ageing error matrix used through the assessment years. It is recommended that this ageing error matrix be updated using data already available. If data is available for the pre-1981 ageing data and if it differs substantially from the post-1981 period, consideration should be given to including that separately.

Size-at-Maturity Data

There are a number of biological properties of the pollock in the Gulf if Alaska that have changed significantly through time. These include the size at maturity, and the weight and length at age. These changes suggest that there are system-wide alterations occurring, and they deserve a great deal of attention as they can have significant effects upon the assessment.

Obviously, if the size and age at maturity alters significantly, this has the potential to have a large effect on estimates of spawning biomass. The unusual and unexpected aspect of the observed changes is that the size and age at first maturity appear to have increased in recent years. As fishing mortality is not very large, compared to natural mortality, and it is possible that natural mortality is changing significantly (see Ecosystem Effects later), then perhaps the expectation that if change were to occur it would be for the fish to mature at a smaller size and younger age was mis-guided. It seems that the size at maturity for males has also changed in the same direction so this phenomenon is broadly applicable across the species. One possibility, discussed in the Ecosystem Effects section, is that there has been some sort of ecological release, either intra-specific competitive release or release from predation pressure.

Currently, a long-term average size at maturity is being used in the assessment, which raises the question of whether there should be more than one vector of such data through time. However, the data is averaged across the Gulf and the changes observed appear localized, so perhaps the impact is smaller than the particular area graphs indicate. If the overall effect is minor, then having more than one size at maturity vector would likely be a greater source of uncertainty than remaining with one vector.

The fact that maturity stages are only visually assessed was highlighted in the discussions. Some of the stages used are very similar to others and may be very brief. It should be a very high priority to conduct histological examinations of the different visual stages to confirm the status of each of the visual guides. While the visual cues have been formalized and photographic guides are provided, there is still a great deal of variability to confuse matters. Fortunately, gonad samples are being taken and stored ready for processing. It is recommended that these gonad samples be processed in an effort to clarify the apparent changes in size at maturity that have occurred.

In terms of the apparent changes in size-at-maturity, because there is less recruitment variation in the Bering Sea, there is potential for comparing this with the Gulf of Alaska; such a comparison could test for the effects of strong year classes on growth and maturity. The non-spawning size-at-maturity is a geographical sub-set of the total area as is the size-at-maturity derived from the spawning ground. The important point is they are different; however, it could be that only the spawning animals move onto the spawning grounds, in which case the size-at-maturity would be biased. Whether the two separate size-at-maturity curves can be combined using a weighting derived from the relative abundance of the different age-classes should be investigated.

Fishery Data

The fishery for Gulf of Alaskan pollock appears to be only a minor player in the dynamics of the stock. This means that the problems with obtaining representative estimates of the size-structure and consequently the age-structure are exacerbated by the small amounts landed at widely dispersed locations.

In terms of observer coverage, there is no requirement for vessels < 60 ft to have observers, but, fortunately, in the Gulf, all vessels were reported to be > 60ft in length overall. These boats then have 30% coverage by observer; however, there is in fact little sampling at sea because of how the fish are processed.

Most sampling is conducted by the observers present in the processing sheds, which provides an adequate coverage of pollock catches but discards require reporting by the at-sea observers. Unfortunately, observer coverage is not random as fishers can elect on which trips they have observers. Because pollock catches tend to be relatively clean, there appears to be a bias towards pollock fishing (as bycatch is less of a problem on such trips). This might have been important because biological data is only taken from an observed trip, so this would also not be random. Once again, however, sampling is probably positively biased towards pollock fishing trips.

Catch rate data are not used in the assessment and, in any case, are only available from observed vessels; the other log-book data is not currently punched. It is possible that this data could be used to corroborate or otherwise the various survey information, but without better information with regard to data quality, this development should be given a low priority relative to the more important work listed elsewhere. It also means that the standardization of the survey vessels suggested earlier may not be possible because of data limitations.

The fishery is required to use mid-water trawls (probably to avoid crab fishery bycatch), which implies that not all of the pollock are necessarily available to the fishery; i.e. the selectivity of the gear would be important. The potential input to the model fit should be investigated via the pseudo-jackknife approach of excluding the different data sources and determining their relative impact on the final fit and levels of uncertainty (width of likelihood profile).

The observer program is multi-species by nature, so it is essential to keep the instructions to observers as general as possible while still getting the sampling completed. Changing these instructions could only be done after extensive consultation as to the practicability of anything new.

Ecosystem Considerations

There have been a number of large and significant ecological changes in the Gulf of Alaska marine system over the last few decades. First, the Steller sea-lion population collapsed in the Gulf of Alaska, and this was a major predator of pollock. Such a decline in a major predator would be expected to permit the pollock population to increase. However, following on from the decline of the Steller sea-lion, the population of another major predator of pollock, the arrowtoothed flounder, appears to have expanded enormously. The steady increase to the average length and weight at age of pollock indicates that some large system change has been happening in a continuous way. This change in biological characteristics is correlated with a large decline in the abundance of pollock. Whether the changes are due to release from predation or release from intra-specific competition, the whole system appears to be well away from equilibrium. This instability implies that any modelling projections are going to be highly uncertain. It appears there may be density-dependent changes occurring within the pollock population and such things are not included in the current assessment model. Fortunately, where the data are available, the observed weight at age is used in the assessment model, so some of the trends are included though not the mechanisms that are driving them.

It is difficult to draw any particular conclusion from such observations; however, there is no doubt that such large shifts away from ecosystem equilibrium will increase the uncertainty in the assessment and in any projections from the assessment of any single species involved. In order to maintain an appreciation of the extent of the possible interactions between the main species of the system, some form of monitoring of the major species is necessary.

The interesting work described in Hollowed *et al*. (2000) strongly suggests that the standard assumption used in many stock assessments, of a constant natural mortality across all ages, does not provide an adequate approximation in the case of the Gulf of Alaska pollock, especially for the younger age classes. In discussions during the review, the suggestion was raised that the model would be extended to include one year old fish. If this suggestion is implemented, then some improvement should be made on the assumption of a constant instantaneous natural mortality of 0.3 for all ages, as is currently the case. No current estimate of natural mortality by age is available and, in any case, with the large changes in the major populations occurring, these values are unlikely to be stable. It may be best, therefore, to investigate the effect on the assessment of various combinations that approximate the parameter values obtained in Hollowed *et al.* (2000). Before the inclusion of the one-year old fish, it would be best to demonstrate that the surveys provide a representative sample of such young animals. The survey results that suggested a strong year class which never eventuated might suggest that natural mortality on the youngest age-class is very high, or, alternatively that only imprecise estimates of the relative abundance of the youngest fish can be produced. If the analyses are consistent, the inclusion of a higher natural mortality for the younger age-classes will lead to an increase in the estimates of spawning biomass.

Stock Assessment Model

Model Structure

The model structure, as defined in the stock assessment document and the AD-Model Builder code is relatively straightforward and standard. It contains idiosyncratic forms for some of the parameters, such as the catchability coefficient, q, which contains what some would refer to as availability, but as long as the particular usage continues to be defined within the documentation then such details are not a problem. The process error treatment of the varying selectivity, as defined in the model, is an excellent solution to annual variations in important model parameters. The model structure is properly and well implemented, and the code is clear and efficiently put together while retaining excellent flexibility.

The possibility of introducing one-year old fish and having a non-equal schedule of natural mortality rates by age was discussed under Ecosystem Considerations. Making such introductions might provide more flexibility and greater realism, but it may not lead to greater precision (it may not lower uncertainty).

Likelihoods

The selection of the specific likelihoods used, log-normal for catches and multinomial for size-structure and age-structure, are standard and classical. The distribution of catches could be graphed to see how well their distribution matches the log-normal distribution, but this is a standard assumption. The other data sources contributing to the likelihoods, such as the egg-production data and survey biomass levels were also standard and no adjustments or changes appear to be required.

One issue that appears to need attention is the partial confounding between age estimation and size frequencies. As discussed in the Numbers-at-Age section, the use of length and weight information to complement the otolith sample when ageing the fish may well be leading to effectively aging by size, at least for the younger, smaller fish, which would be expected to form discrete modal size groups. What this implies is that even though the length data modelled is independent of the lengths used to help in the generation of the ageing data, the data are not strictly independent. To what degree they are inter-dependent is not a simple question nor easy to determine; it would depend partly on strong year classes entering the fishery and on how long it takes for modal length classes to merge as the fish get older and bigger.

At present, by using likelihoods for both the catch-at-age data and the catch-at-length data, because they are not strictly independent, the assessment model would be under-estimating the overall uncertainty. The best solution would be to generate the catch-at-age data independently from the length or weight data. Alternatively, the younger size-classes could be given less weight in the fitting process. One, possibly weak, test of the confounding between ageing data and length data is to determine whether each were providing the same signal to the model (as would be expected if they were really confounded).

Current Management

The present management is based upon allocating a set proportion of the spawning biomass to the fishery. A question explicitly asked of the review was whether the current management approach is precautionary, given current stock status and trends? The current spawning stock size appears to be smaller than previously experienced. However, there also appears to be a strong year class entering the fishery from 1999 and possibly in 2000. The biomass surveys used in the assessment tend to be conservative and are likely to be under-estimating the biomass available to the fishery. All of these estimates are uncertain but, on top of this, the ecosystem of which the pollock are a part is showing clear indications of great instability.

Overall, the Gulf of Alaska pollock are extremely difficult to assess. However, the management strategy adopted in the Tier 3 schedule, used with the pollock fishery, appear to be a reasonable response to the present situation. The fishery appears to form only a relatively minor part of the dynamics of this system and the estimates of spawning biomass appear, on balance, to be conservative. Under these conditions, the present management options should be sufficiently robust. If the estimates of spawning stock biomass continue to decline, the difficulty will become one of deciding whether to believe the estimates if they fall just below the threshold of 20% unfished spawning biomass. Such estimates will be uncertain but the management response appears to be firmly in place. It would be prudent to state that the assessment is currently the best available information and encourage the management agency to decide whether it is necessary to add a risk clause to the management (e.g. if the upper 75 percentile of the likelihood profile around the spawning biomass estimate is below the threshold then the management trigger is activated). Some form of risk assessment is required rather than a simple threshold trigger for the management response.

Conclusions/Recommendations

The Gulf of Alaska pollock assessment and management appears to be relatively well developed with few un-explored options. Because of the variable ecosystem of which it is a part, the pollock fishery is an exceptionally difficult species for which to generate a highly certain assessment. Many of the uncertainties lead to a conservative assessment (e.g. the survey estimates are often biased low, the constant natural mortality for all ages biases the spawning biomass low, the analysis of Eulachon in the EIT biases estimates of pollock low, etc), which confirms that the current management is precautionary. It might be an issue that with the fishery being such a small part of the dynamics that the fishery management might be overly precautionary.

Overall, the assessment and surveys that contribute to it are well structured. We have made a number of suggestions that might lead to improved estimates of biomass or at least of the uncertainty associated with the assessment. These suggestions are more directions for exploration than fixed ideas.

The current management is certainly precautionary in the context of the present assessment to stock trends and its associated uncertainties. It is recommended that rather than using an abrupt decision rule based on an uncertain assessment some form of risk assessment be developed instead that suggests management responses in the light of the varying likelihood of different events occurring (some form of decision table be implemented). If the median assessment of the spawning stock biomass falls below a limit reference point, this does not necessarily entail that the stock is definitely below that limit. Some way of accounting for the uncertainty inherent in the assessment should be included in the management decision process.

Through the text of the review document, a number of suggestions have been made with regard to potential further work, alternative approaches, clarifications and sampling strategy. To reiterate the major possibilities:

- If formal trials to compare the fishing power of the different vessels used in the bottom trawl survey are not available, then classical standardization of fishing data using GLMs, including vessel as a factor, could provide an indication of their relative efficiency (if they have fished in similar areas using similar gear, and the data are available).

- If it is possible to correct the catch rates of the pre-1996 bottom trawl surveys for distance towed, using the calibration generated from using the custom bottom sensing gear, then this should be done.

- The fact that the time-series of bottom trawl biomass estimates is really made up of two time series (with either a step change between them or a slower transition between the pre-1996 and the following surveys) could be recognized in the assessment modelling. This would reduce the number of data points available for the estimation of each set of the relevant parameters but would more closely reflect the reality of the situation.

- Strong consideration should be given to moving the surveys towards using fixed stations, which would have many advantages, especially in the Gulf of Alaska where there are significant areas of untrawlable ground.

- The current swept area estimates of biomass include the untrawlable ground area. Without knowledge of whether pollock are less or more dense over these regions, this is possible a large source of uncertainty. The impact of including and excluding this area when conducting the calculations should be investigated.

- It was suggested during the review that the use of a trawl spreader limitation device would standardize the spread ratio of the trawl mouth, which would increase the validity of treating the trawl abundance estimate as an absolute estimate of abundance. However, to make such a change would instigate another time-series of abundance indicators; so if this is deemed desirable, then it should be implemented along with other changes that have been recommended (such as going to fixed stations). In the meantime, the bottom trawl abundance estimates should be considered as giving an index of relative abundance.

- For the EIT survey it was recommended that the current strategy of expanding the survey to cover spawning grounds beyond the Shelikoff Strait be continued and enlarged. Similarly, if an EIT survey could be conducted in the summer to corroborate or otherwise the bottom trawl surveys this would also prove advantageous.

- There was discussion on how to obtain improved estimate of target strength, of comparing the analysis both with Eulachon and with Eulachon omitted from consideration (because of not having a swim bladder). Both of these would improve the survey estimates.

- Putting in place a strategy for including the dead zone at the sea-floor in the EIT surveys would also improve, and likely increase, the biomass estimates.

- Finally, with respect to the trawling, Dr Godo suggested that a forward looking sonar would provide some degree of knowledge about the spread of fish in front of the trawl and what proportion are being sampled by the net. This could be a highly significant influence on the biomass estimates and should be given a reasonably high priority.

- While there are limited opportunities for modifying the Alaska Department of Fish and Game surveys it was recommended that by systematically excluding the different data sources in a pseudo-jackknife procedure would enable a determination of whether the different data sources are consistent with each other and which are contributing to the overall uncertainty.

- A high priority should be given to investigating the effect of providing information on the length and weight of fish from which otoliths are taken for ageing. The methods used to describe ageing error are standard, but the possibility of introducing bias or of confounding length measurements with age estimates (given strong year classes and relatively clear size modes) means that some doubt exists over the use of both catch-at-age and catch-at-length data in the same assessment.

- At the very least, the ageing error matrix should be updated with the latest estimates of reading error using already available data.

- Various biological parameters have been found to vary through time including the size-at-maturity, the length-at-age, and the weight-at-age. As maturity plays a large part in determining the performance measure for this fishery, it is essential that this variability be either explained or corrected. A high priority should be given to processing samples of gonads collected at the different visually determined maturity stages to provide an objective statement of exactly what ovary conditions indicate sexual maturity.

- Investigations could proceed with the stock assessment model to determine the impact of alternative natural mortality schedules for the younger age classes. This, combined with inclusion of one-year old fish, might introduce significant changes to the outcomes from the model. This, combined with the investigation of the relative consistency of the various data sources and their relative contribution to the overall uncertainty (as perhaps expressed by the width of the likelihood profiles), would enable a more defensible assessment to be written.

- If the downward trend in the spawning stock biomass continues, then the management rules will recommend that either drastic cuts to catches must occur or fishing ceases entirely. It would be useful to clarify, and possibly expand the details of, these rules to the Industry and the management advisory bodies so that if this stage eventuates there will be less argument.

Appendix I

Bibliography

Alaska Fisheries Science Center (1998) A Guide to the Preparation of Alaska Groundfish SAFE Report Chapters. Unpublished MS, 5 p.

Anon. (1990) Report from the Workshop on Ageing Methodology of Walleye pollock (*Theragra chalcogramma*), Held at the Sea Fisheries Institute, Gdynia, Poland, September 10-14, 1990.

Anon. (1998) Report from the Second Workshop on Ageing Methodology of Walleye pollock (*Theragra chalcogramma*), Held at the Alaska Fisheries Science Center, Seattle, Washington, U.S.A. March 17-20, 1998.

Anon. (2001) Cruise Results: Cruise 2001-01 Chartered F/V's *Morning Star*, and *Vesteraalen* 2001 Gulf of Alaska Biennial Groundfish Assessment survey May 17 – July 25, 2001. 11p.

Anon. (2002) Introduction to Stock assessment and fishery evaluation report for the groundfish resources of the Gulf of Alaska. Prepared by the Gulf of Alaska Groundfish Plan Team, North Pacific Fishery Management Council, P.O. Box 103136, Anchorage, AK 99510. North Pacific Fisheries Management Council, Anchorage, AK.

Bailey, K.M., Quinn II, T.J., Bentzen, P., and W.S. Grant (1999) Population structure and dynamics of walleye pollock, *Theragra chalcogramma*. *Advances in Marine Biology* **37**: 179-255.

Bailey, K.M. (2000) Shifting control of recruitment of walleye pollock *Theragra chalcogramma* after a major climatic and ecosystem change. *Marine Ecology Progress Series* **198**:215-224.

Britt., L.L., and M.H. Martin. (2001) Data report: 1999 Gulf of Alaska bottom trawl survey. NOAA Technical Memorandum NMFS-AFAC-121.

Dorn, M., Barbeaux, S., Guttormsen, M., Megrey, B., Hollowed, A., Brown, E., and K. Spalinger (2002) Assessment of Walleye pollock in the Gulf of Alaska. In *Stock Assessment and Fishery Evaluation Report for Groundfish Resources of the Gulf of Alaska.* Prepared by the Gulf of Alaska Groundfish Plan Team, North Pacific Fishery Management Council, P.O. Box 103136, Anchorage, AK 99510. North Pacific Fisheries Management Council, Anchorage, AK.

Guttormsen, M., Wilson, C., and S. Stienssen (2002) Echo integration-trawl survey results for walleye pollock (*Theragra chalcogramma*) in the Gulf of Alaska during February and March 2002. Appendix D In *Stock Assessment and Fishery Evaluation Report for Groundfish Resources of the Gulf of Alaska*. Prepared by the Gulf of Alaska Groundfish Plan Team, North Pacific Fishery Management Council, P.O. Box 103136, Anchorage, AK 99510. North Pacific Fisheries Management Council, Anchorage, AK.

Hollowed, A.B., Ianelli, J.N., and P. Livingston (2000) Including predation mortality in stock assessments: a case study for Gulf of Alaska pollock. *ICES Journal of Marine Science* **57**:279-293.

Kimura, D.K. and J.J. Lyons (1991) Between reader bias and variability in the age-determination process. *Fishery Bulletin* **89**: 53-60.

Kimura, D.K., Lyons, J.J., MacLellan, S.E., and B.J. Goetz (1992) Effects of year-class strength on age-determination. *Australian Journal of Marine and Freshwater Research* **43**: 1221-1228.

Kimura, D.K. A brief history of age-determination of Walleye pollock (*Theragra chalcogramma*) at the Alaska Fisheries Science Center. www.afsc.noaa.gov/refm/age/brief_history.htm

Meuter, F.J. and B.L. Norcross (2002) Spatial and temporal patterns in the demersal fish community on the shelf and upper slope regions of the Gulf of Alaska. *Fishery Bulletin* **100**:559-581.

NPGOP (2003) *North Pacific Groundfish Observer Program Overview 2001.* Resource Ecology & Fisheries Management Division, Alaska Fisheries Science Center, national Marine Fish Service, Seattle, Washington. 44p.

Olsen, J.B., Merkouris, S.E., and J.E. Seeb (2002) An examination of spatial and temporal genetic variation in walleye pollock (*Theragra chalcogramma*) using allozyme, mitochondrial DNA, and microsatellite data. *Fishery Bulletin* **100**:752-764.

Shima, M., Hollowed, A.B., and VanBlaricom, G.R. (2002) Changes over time in the spatial distribution of walleye pollock (*Theragra chalcogramma*) in the Gulf of Alaska, 1984 – 1996. *Fishery Bulletin* **100**: 307-323.

Spalinger, K. (2003) Bottom Trawl Survey of Crab and Groundfish: Kodiak, Chignik, and South Peninsula Management Areas, 2002. *Regional Information Report* **No 4K03-32** Alaska Department of Fish and Game, Division of Commercial Fisheries, 211 Mission Road, Kodiak, Alaska 99615.

von Szalay, P.G. and E. Brown (2001) Trawl comparisons of fishing power differences and their applicability to National Marine Fisheries Service and Alaska Department of Fish and Game Trawl Survey Gear. *Alaska Fishery Research Bulletin* **8**: 85-95.

von Szalay, P. G. (In review) The feasibility of reducing the variance of fish biomass estimates by integrating catch per unit effort data from two demersal trawl survey in the Gulf of Alaska. *Alaska Fishery Research Bulletin* **10**

The Plan Team for the Groundfish Fisheries of the Gulf of Alaska (2002) Appendix B. Stock Assessment and Fishery Evaluation Report for the Groundfish Resources of the Gulf of Alaska. North Pacific Fishery Management Council, 605 W 4th Avenue, Suite 306, Anchorage, AK 99501

Wilson, C.D., Hollowed, A.B., Shima, M., Walline, P. and S. Stienessen (2003) Interactions between commercial fishing and Walleye pollock *Alaska Fishery Research Bulletin* **10**: 61-77.

Worton, C. (2000) Bottom trawl survey of crab and groundfish: Kodiak Island, Chignik, and South Peninsula areas, 1998. *Regional Information Report No. 4k00-20*. Alaska Department of Fish and Game, Division of Commercial Fisheries, 211 Mission Road, Kodiak AK 99615 (more recent reports available from ADFG?)

Appendix II

STATEMENT OF WORK

Consulting Agreement Between The University of Miami and Dr. Malcolm Haddon

General

The Alaska Fisheries Science Center (AFSC) requests review of the Gulf of Alaska walleye pollock (*Theragra chalcogramma*) assessment. pollock in the Gulf of Alaska have declined in abundance, and in 2002 dropped to a 30 year low, despite being managed conservatively under North Pacific Fishery Management Council harvest policies. Both acoustic and trawl surveys are used to assess pollock abundance, but interpretation of recent survey results has been problematic due to contradictory trends and potential changes in stock distribution. Reliable assessment of Gulf of Alaska pollock stock status is important both for council management of the fishery and for evaluating fishery impacts on endangered Steller Sea Lions, for which pollock are important prey. These factors create a compelling need for independent peer review of the Gulf of Alaska pollock assessment.

The assessment review will require two consultants, 1) a stock assessment expert, and 2) an expert on survey methodology for fish resources (both acoustic and trawl surveys). Consultant 1 should to be thoroughly familiar with various subject areas involved in stock assessment, including population dynamics, separable age-structured models, harvest strategies, the AD Model Builder programming language, and have experience conducting stock assessments for fisheries management. Consultant 2 should be thoroughly familiar with survey methodology (both bottom trawl and acoustic), its application in stock assessments, and have experience planning and conducting resource assessment surveys. The consultants will travel to Seattle, Washington, to discuss the stock assessment with the lead analyst, the chief survey scientist, and other scientists at the Alaska Fishery Science Center involved in the Gulf of Alaska pollock assessment. The report generated by the consultant(s) should include:

a. The strengths and weaknesses of the Gulf of Alaska pollock assessment and harvest strategy;
b. Recommendations for improvement to the assessment model and harvest strategy;
c. Strengths and weaknesses of current survey for assessing Gulf of Alaska pollock abundance trends;
d. Recommendations for survey improvements to better assess Gulf of Alaska pollock;
e. Suggested research priorities to improve the stock assessment.

AFSC will provide copies of stock assessment documents, AD Model Builder code for the stock assessment model, survey reports, and other pertinent literature.

Specific issues that should be addressed in the review are the following:

a. The appropriateness and statistical rigor of the assessment approach.
b. Adequacy of existing surveys to monitor abundance trends of GOA pollock.
c. Whether the stock assessment takes into account major uncertainties in data and assumptions.

d. Whether the current management approach is precautionary, given current stock status and trend.

Specific

1. Read and become familiar with the relevant documents provided to the consultant;
2. Discuss the stock assessment with the lead assessment scientist and survey scientists in Seattle, Washington, from August 4 to August 8, 2003;
3. No later than August 22, 2003, submit a written report[1] consisting of the findings, analysis, and conclusions, addressed to the "University of Miami Independent System for Peer Review," and sent to Dr. David Die, via email to ddie@rsmas.miami.edu and to Mr. Manoj Shivlani, via email to mshivlani@rsmas.miami.edu.

[1] The written report will undergo an internal CIE review before it is considered final. After completion, the CIE will create a PDF version of the written report that will be submitted to NMFS and the consultant.

ANNEX I: REPORT GENERATION AND PROCEDURAL ITEMS

1. The report should be prefaced with an executive summary of findings and/or recommendations.

2. The main body of the report should consist of a background, description of review activities, summary of findings, and conclusions/recommendations.

3. The report should also include as separate appendices the bibliography of materials provided by the Center for Independent Experts and the center and a copy of the statement of work.

Please refer to the following website for additional information on report generation: http://www.rsmas.miami.edu/groups/cie