Summary Review Report
for CIE Review of Alaska flatfish stock assessments, June 11-15, 2007

Prepared by:
Cynthia M. Jones, Ph. D.

Prepared for
Center for Independent Experts (NTVI)

July 2007

Table of Contents

I. Executive summary of findings

The Alaska Fisheries Science Center (AFSC) requested a peer review of the Gulf of Alaska and Bering Sea/Aleutian Islands flatfish stock assessments by the Center for Independent Experts (CIE). The review covered the assessments of arrowtooth flounder (*Atheresthes stomias*) and rex sole (*Errex zachirus*) in the Gulf of Alaska, and Greenland turbot (*Reinhardtius hippoglossoides*), northern rock sole (*Lepidopsetta polyxystra*) and yellowfin sole (*Limanda aspera)* as examples of the types of approaches and methodologies that are conducted by the BSAI and GOA Groundfish Plan Teams. These stock assessments have never undergone outside review and such review is timely given that the North Pacific Fisheries Management Council is likely to pass new amendments for the flatfish fisheries. Three reviewers, Drs. Din Chen, Paul Medley and Graham Pilling, constituted the Review Panel which was convened during June 11-14th 2007 in Seattle, Washington, at the NMFS Alaska Fisheries Science Center. The reviewers were given three Terms of Reference that included:

> TOR 1. Modeling efforts for Bering Sea/Aleutian Islands (BSAI) and Gulf of Alaska (GOA) flatfish assessments and harvest recommendations. Specifically, the review shall evaluate: a) The analysts' use of fishery dependent and fishery independent data sources in the assessments; b) Gaps or inconsistencies in the population dynamics modeling methodology or logic; c) If uncertainties in assessment model results are appropriately applied to management advice; and d) Whether the assessments provide the best available science.
>
> TOR 2. The effort to incorporate ecosystem indicators and shifts in states of nature in the assessments. These include modeling survey catchability with annual bottom water temperature and using the Ocean Surface Current Simulation Models (OSCURS) to define putative oceanic productivity regimes.
>
> TOR 3. The harvest control rules adopted for Bering Sea yellowfin sole and northern rock sole (where a stock-recruitment model and F_{MSY} quantities are estimated) compared to other flatfish stocks where proxy values are used. Specifically, comments on the trade-offs between the different approaches are required.

There was consensus that the modeling for BSAI and GOA flatfish assessments and harvest recommendations were appropriate given the available data and provided the best scientific advice.

Under TOR 1, the reviewers all felt that the fisheries-independent trawl survey provided reasonable estimates of abundance, given that it was not designed for that purpose, and also that the sablefish longline survey also provided a reasonable index for Greenland turbot. They were concerned that budget constraints might imperil data collection, specifically on the Bering Sea slope.

For fishery-dependent data, they noted that there was a long time series which is supplemented well by observer coverage. All three reviewers noted that catch data was reliable for flatfish, with the exception of Greenland turbot. Because of the observer program, they felt that discards and by-catch were also well recorded. Their concern with the fishery-dependent data was based on the migration of Greenland turbot between U.S. and Russian waters and the potential for evaluating only a portion of the stock, which would lead to a miss-estimation of stock biomass.

There was also consensus among the reviewers that the models being used were appropriate. These models include AD Model Builder and SS2. However, the spatial extent and potential migration of Greenland turbot made these models less reliable for that species. The reviewers also were concerned with the use of a single sex model for yellowfin sole. Because there are considerable sexual differences in size that can result in mortality differences, the reviewers thought that a two-sex model would be more appropriate. The reviewers agreed that uncertainty was handled appropriately in the models because assessment scientists are using AD Model Builder with MCMC bootstrapping. However, the reviewers also thought that there need to be further work in this area and more formal procedures to include uncertainty in recommendations. Altogether, the reviewers stated that the assessments generally provided the best available science and Dr Pilling noted that “the work performed is impressive.”

Under TOR 2, the reviewers discussed the various modeling approaches that could support ecosystem-based management for BSAI and GOA flatfish. They specifically addressed the application of Ecopath with Ecosim (EwE); modeling of bottom temperature effects on selectivity, distribution, growth, and fecundity; Ocean Surface Current Simulation Models (OSCURS). There was consensus among the reviewers that the use of EwE was worthwhile because it pointed out the importance of the flatfish community in the ecosystem. However, they also noted that they expected that there would be limitations with EwE because flatfish are data poor and that other models should be developed, e.g. “Atlantis”.

The reviewers did have more comments on the models that related temperature to factors such as selectivity, but generally agreed that this modeling was worthwhile. They suggested various approaches to improve these models, including 1) examining the indirect effects of temperature on the distribution of fishing fleets as relates to stock abundance, and 2) developing formulations that were related to metabolic theory.

All three reviewers commented on the value of using OSCURS to model productivity. The liked the fact that the model had been validated and saw the larval dispersal model as an interesting scientific study. However, they also noted the inherent difficulties in predicting the abundance of adults from early life stages, and thought that modeling post-settlement stages would provide greater predictive value. The reviewers comments were rich in detail and their reports are attached as appendices.

For TOR3, the reviewers were again largely in agreement that the tier system should be simplified. The North Pacific Fisheries Management Council uses a 6-tier system where a stock is assigned to a tier based on the extent and quality of its data. The tier system was designed to allow greater exploitation as data on the stocks improves, and is a precautionary approach to management. Several reviewers commented on a lack of documentation defining the system clearly (but see Goodman et al 2002) and also thought that and that there was no substantive research to support the use of the definitions of F_{OLF} or F_{ABC} within each tier. The reviewers also agreed that BSAI yellowfin and rock soles had very reliable data and should be moved to tier 1. The other concern was the use of 0.2 as the value for natural mortality (M), The reviewers found evidence from other flatfish stocks that M could be lower than this and that species that used this assumed value should actually be classified as tier 6 not tier 4 or 5.

The review resulted in 32 recommendations. Of these two were endorsed by all three reviewers and include 1) that more research and effort should be made in the development of multispecies/ecosystem modeling and management, and 2) the 6-tier system for harvest controls should be simplified for flatfish. Besides these, there were 9 more recommendations endorsed by two of the reviewers. Care should be taken in interpreting the level of endorsement, given that each reviewer has their own writing style and if they were presented with the final list, they could all agree.

II. Summary Report of the Flatfish Review Panel, 2007

Background

The purpose of the review was stated in the Statement of Work as follows. "The Alaska Fisheries Science Center (AFSC) requested a peer review of the Gulf of Alaska and Bering Sea/Aleutian Islands flatfish stock assessments by the Center for Independent Experts (CIE). Flatfish fisheries in Alaska are valuable and are likely to receive increased interest from commercial fishers in the next decade. Most flatfish populations are currently at stable or high levels of abundance. The North Pacific Fisheries Management Council (Council) is likely to pass amendments to rationalize (establish quota shares) groundfish fisheries in the near future. These amendments will allow for more flexibility in the time and areas fished which may result in the development of renewed interest in flatfish fisheries. Thus the AFSC desires an independent peer review of these stocks to assess the quality of the assessments and to determine whether the Council is being provided with the best available information and analysis."

There are currently 21 flatfish species managed under the Bering Sea and Aleutian Islands Fishery Management Plan, six of which have age-structured stock assessments. In the Gulf of Alaska there are ten managed species, four of which have age-structured assessments. These assessments have never undergone a CIE review. The main stocks assessments covered by the CIE review were arrowtooth flounder (*Atheresthes stomias*) and rex sole (*Errex zachirus*) in the Gulf of Alaska, and Greenland turbot (*Reinhardtius hippoglossoides*), northern rock sole (*Lepidopsetta polyxystra*) and yellowfin sole (*Limanda aspera)*. For each of these species a "stock assessment and fishery evaluation (SAFE) report is prepared and reviewed annually for each fishery management plan (FMP)" and were produced by the BSAI and GOA Groundfish Plan Teams. These reports "include a recommendation for the overfishing level (OFL) and acceptable biological catch (ABC) for each stock and stock complex managed under the FMP." These individual stock assessments were not evaluated in detail separately, but rather reviewed generally so as to comment on the general strategy and methods used in the assessment and whether this provides the best available science.

Description of review activities

My role was to provide a summary report to the CIE based on reading and becoming familiar with the relevant documents provided to the reviewers, and by reading the draft reports from the individual CIE reviewers. I was not present at the CIE panel meeting at the NMFS Alaska Fisheries Science Center during June 11-14th 2007. Instead, I received documents relevant to the summary report by email during July 2007. I obtained three CIE panelist reports written by Drs. Din Chen, Paul Medley and Graham Pilling. In addition to these reports my reading included the ten age-structured assessments, the Goodman et al (2002) report on the PFMC harvest control strategy, and supporting materials from http://www.afsc.noaa.gov/REFM/Stocks/assessments.htm and ftp://ftp.afsc.noaa.gov/afsc/public/flatfish/flats.html.

The three reviewers, Drs. Din Chen, Paul Medley and Graham Pilling, the Review Panel held in Seattle, Washington, at the NMFS Alaska Fisheries Science Center. The Flatfish Review Panel was convened during June 11-14th 2007. The agenda and participants are listed in the appendices within the individual reviewer reports.

Summary of findings

All three reviewers commented that the modeling efforts for BSAI and GOA flatfish assessments and harvest recommendations were appropriate given the available data.

TOR 1. Modeling efforts for Bering Sea/Aleutian Islands and Gulf of Alaska flatfish assessments and harvest recommendations.

Specifically, the review shall evaluate:

a. The analysts' use of fishery dependent and fishery independent data sources in the assessments;

Fishery independent data

Fishery-independent surveys are used to determine relative abundance of flatfish and include trawl and longline surveys. The trawl survey has been used as an index of flatfish abundance even though it was not designed for that purpose. Thus, Dr. Medley noted that he would normally be concerned that it "may miss significant biomass, be noisy and subject to climatic and other effects". However, all three reviewers felt that the trawl survey provided reasonable estimates nonetheless. Additionally, sablefish longline surveys were used as an index in the Greenland turbot assessment, but this did not elicit concern. Each reviewer emphasized different aspects of the fishery-independent data that each thought was well done.

In terms of strengths of the fishery-independent data, they mentioned that age estimation was reliable, that there was a good time series of data for both the BSAI and GOA, that the likelihoods used in the surveys were generally appropriate. Given that aggregated catches are evaluated, zero catches are unlikely, but Dr. Medley commented that this should be checked. In contrast for the longline data, assessment scientists might consider alternatives such as the binomial, Poisson negative binomial or beta-binomial. The reviewers commented that the work on temperature and trawl-gear herding showed great promise and should be continued. Dr. Medley and Dr. Pilling specifically noted that "research on trawl efficiency, the herding effect and standardisation using bottom temperature and types of escapement from the survey net should significantly improve the survey indices." The use of the exponential model to describe the relation between temperature and catchability was a good choice. The use of cameras was also lauded because it provides direct observation of behaviour as flatfish encounter the gear.

In terms of weaknesses of fishery-independent data, the reviewers mentioned specifically that they were concerned that budget constraints imperiled data collection, specifically in the reduction of trawl surveys on the Bering Sea slope. One suggested that NMFS should “look at the time series of information from the commercial fishery and in particular the time series of data from the commercial observer programme, as an alternative index of abundance.”

Fishery dependent data

Fishery-dependent data consists of a long time series of catch data (1960-2005), and an observers program that provides 100% coverage of vessels 125 ft and greater, 30% for vessels between 60 and 125 ft, and 0% for vessels under 60ft. Observers also record catch composition, bycatch, effort, location, and obtain biological samples and metrics.

The strengths of the fishery-dependent data are the availability of a long time series supplemented by observer coverage. All three reviewers noted that catch data was reliable and well estimated for flatfish, with their only concerns being for Greenland turbot. They stated that the “sampling methods and coverage are appropriate and well documented” and that there was excellent training and quality control in the observer program. Because of the observer program, the reviewers believed that the discards and by-catch were well recorded. Dr. Medley expressed mild concern that these data would be influenced by an observer effect, which could be evaluated, but did not believe that this threatened the validity of the assessments.

The weakness of the fishery-dependent data was identified as the catch data for Greenland turbot, which was considered “the least reliable stock assessment.” In part, this concern involved problems with potential miss-estimation of stock biomass because of migration between US and Russian waters, which will be discussed in the next section. However, Dr. Medley noted his concern that there was less age data for Greenland turbot than for other flatfish species, “so there is greater reliance on size composition data from which age is inferred”, leading to more uncertainty. Dr. Medley also noted that, although there was a long catch series, that there was a gap in coincident survey coverage, which prevented a direct estimate of stock and recruitment. When Greenland turbot was under heavy exploitation it was covered only by catch data, and the survey data commenced only after high exploitation. Dr. Pilling commented that “It is difficult to identify the impact of these issues on assessment results (including the stock-recruitment data, ABC and OFL calculations).” Dr. Pilling raised his concerns that cessation of the fishery-independent survey would also be detrimental to the stock assessment for this species because potential bias in the catch could not be determined.

Dr. Paul Medley specifically checked the results of the current SS2 model for potential bias for Greenland turbot by applying a simple biomass dynamics model (Schaefer), which was fitted to the catch and survey data. He was reassured that “the biomass dynamics model does produce approximately the same trajectory of the biomass in recent years as the catch-at-age model.” Hence, the Schaeffer model could be used to check estimates of biomass early in the time series.

Migration issues

In part, the validity of the stock assessments rely on the assumption that catch and survey data represent the entire stock. The reviewers stated that this was a viable assumption in the case of both BSAI and GOA flatfish stocks which are generally assumed to have little migration. This assumption was further supported by ad hoc tagging studies that showed little migration. Dr. Pilling stated that "Anecdotal evidence suggests that the catches of flatfish in neighbouring waters are minimal, based on the limited abundance levels in US waters neighbouring Canada, and the small seabed shelf area in the Russian zone." However, the reviewers were concerned that this assumption might not hold for Greenland turbot. In the case of Greenland turbot, the species is found on the slope in sufficient abundance to indicate that the stock extends into the Russian zone. Additionally, limited tagging data does suggest movement occurs between the U.S. and Russian waters. It is assumed now that fish in U.S. water represent 75% of the total stock, and this assumption is used in the SS2 model. At least in the case of Greenland turbot, more studies are warranted to clarify stock delineation and movement.

Life History Data

Life history data was mostly discussed in context to issues such as sex-specific density-dependent growth as a stock assessment model input. However the reviewers did comment specifically to say that the "assumed age-structure population dynamics are consistent with the current understanding and knowledge of the flatfish biology" and "Natural mortality estimate is a notoriously difficult parameter to estimate. For the flatfish stocks, this M is even problematic and used 0.2 for most of the stock. This is especially problematic for stock in tier 5."

b. Gaps or inconsistencies in the population dynamics modeling methodology or logic;

Assessment models

The assessment models for the flatfish species use AD Model Builder as a generic model of age-structured population dynamics models, and Stock Synthesis 2 was used when simpler or alternate methods were needed. All three reviewers commented that these models were appropriate for the data. These models do not consider migration of spatial structure. For now, they are appropriate but when more data becomes available on migration (especially for Greenland turbot) the models should be modified if necessary.

The strengths of the stock assessment models are that the use of the Ricker stock-recruitment relationship provides a more precautionary assessment, "at least at the high stock sizes currently estimated." Although the Beverton-Holt stock-recruitment relationship is more commonly used when there is no biological evidence to justify density responses at high stock size, nonetheless the Ricker provides a better fit to the data. Dr. Pilling also commented that "impact of stock-recruitment relationship model uncertainty on the performance of [stock] management should be examined through simulation testing…"

The weakness of the stock assessment models concern the use of a single sex model for yellowfin sole, the lack of sensitivity investigation for Bayesian priors, and the potential problem of parameter confounding. First, because of large growth differences between sexes of yellowfin sole, a two-sex model is preferable to the single-sex model used now. Although the single sex model may be reasonable under current conditions, it "has the potential to cause a bias where selectivity is sex-preferential (e.g. towards females due to faster growth and greater size)." Second, Dr.

Chen, whose expertise is Bayesian models, stated that "it should be prudent for any Bayesian modeling for sensitivity investigation for the appropriate choices of the prior distributions." Finally, all three reviewers commented in detail about their concern with parameter confounding and local optimization, when so many parameters were being estimated. Dr. Pilling suggested that "The inverse Hessian matrix and MCMC evaluation of the posterior are both routinely examined for confounding between parameters or poor estimation." They believed that the problem will become greater in the future.

Other weaknesses were commented upon by Dr. Pilling, who was quite thorough. He had six concerns with the use of the stock assessment models. First, he was concerned that there were no formal retrospective analyses "to identify persistent assessment biases and uncertainties." Second, he had "some uncertainty over whether the level of fishing/total mortality is being estimated well", but also felt that the dearth of tag returns and age structure did support the estimation of low fishing mortality that was found with the model. Third, he was more concerned about the uncertainty for the natural mortality that was selected for the flatfish models, especially because "differential natural mortality at size (rather than age) might explain some of the sexual dimorphism seen in the flatfish (e.g. Gulf of Alaska arrowtooth flounder)", Fourth, he thought that the "impact of using a mean weight-at-age or –at-length for the plus group" should be monitored. Fifth, he stated that "some notable changes have occurred in length-at-age" of GOA rex sole "between 1984 and 1990, it would be useful to obtain additional age information for more recent years to ensure that no biases occur in the assessment as a result of any unidentified trends…". Sixth, he felt that the model constraints should be relaxed for yellowfin sole in an attempt to fit older data for which whole otoliths were used in ageing and where under-ageing was possible. This reviewer also addressed additional problems with the Greenland turbot assessment. Specifically, he was concerned that "the projection model uses age-specific future selectivity, whereas the assessment model is based upon size-selectivity (J. Ianelli, pers. comm.)." Among the two models "[t]he current generic model aims to allow the linking of bycatch information (which represents the current extent of the fisheries management system's consideration of the multispecies nature of fisheries) with single species assessments. The progression of this multispecies projection model is recommended."

c. If uncertainties in assessment model results are appropriately applied to management advice

The reviewers agreed that the flatfish assessments were appropriate and captured uncertainty by using AD Model Builder with MCMC and bootstrapping. The degree of uncertainty is not used directly to supply confidence intervals, but is used to help decide which management tier to which a stock is assigned although it is reported at the assessment stage. "Alternative ABC and OFL levels may be provided by assessment authors based upon different data sources etc. (e.g. Gulf of Alaska rex sole). Furthermore, the level of uncertainty in the assessment defines the management tier into which the stock is placed and hence the harvest rule used." The reviewers also agreed that there was a need to have a more formal representation of uncertainty in assessment outputs. In projection models "considerations of uncertainty are limited to recruitment variation." "A better consideration of uncertainty could be included through the inclusion of uncertainty in current year numbers-at-age, for example."

d. Whether the assessments provide the best available science.

All reviewers noted that the stock assessments generally provided the best available science and one reviewer noted that "the work performed is impressive." There assessments are supported further by evidence that mortality is low based on age structure, tag returns, and catch curve analysis performed during the review by Dr. Medley. Dr. Medley suggested that alternate assessment methods (catch curves, VPA) should be used periodically in support of the stock assessments.

TOR 2. The effort to incorporate ecosystem indicators and shifts in states of nature in the assessments. These include modeling survey catchability with annual bottom water temperature and using the Ocean Surface Current Simulation Models (OSCURS) to define putative oceanic productivity regimes.

Ecosystem-based modeling approaches have included models such as Ecopath-with-Ecosim (EwE), the Ocean Surface Current Simulation Models (OSCURS), and "modeling survey catchability with annual bottom water temperature and using to define putative oceanic productivity regimes". As Dr. Medley pointed out, this modeling has four goals: 1) Improve estimates of natural mortality by accounting for predation, 2) Explain changes in abundance patterns, particularly for arrowtooth flounder, that have been observed in the shelf ecosystem, 3) Understand changes in ecosystem patterns that result from climate change, and 4) Develop multispecies and ecosystem methods that estimate optimum yields. The reviewer reports were less uniform for this TOR, as each had a slightly different focus. This was somewhat more difficult to synthesize than when the points addressed were similar, as they were in the other TORs.

The reviewers all commented positively about the use of EwE and generally thought that this modeling approach was worthwhile because it pointed out the importance of the flatfish community in the ecosystem. However, they also noted that the model was better suited to data-rich situations rather than Alaskan flatfish which were considered to be data poor. It was felt that the models would "eventually fall short of what is required, and it will be necessary to move to a more detailed process model (e.g. "Atlantis" and "In Vitro" modeling approaches used in Australia)." Dr. Pilling pointed out that the models would be improved with the continuing collection of stomachs from the observer program. However, they viewed the model as largely qualitative and that multispecies models should be developed further to provide better guidance for fisheries management.

Dr. Chen and Dr. Pilling commented on the work being done to model flatfish distributions in relation to temperature changes using survey data, stating this modeling approach was a strength of the program and that the assessment scientists had shown "a strong statistical relationship between fish distribution and cold pool distribution have been found for several flatfish species." Dr. Pilling suggested that the VMS data be used to examine "indirect impacts of temperature on the distribution of fishing fleets in a similar way." He stated that the "degree of overlap between fleet and stock abundance under scenarios of future temperature change may give some idea of the relative levels of bycatch and discarding of flatfish under these different climate hypotheses."

Further, these two reviewers also noted the evidence that incorporating temperature effects on survey catchability was also worthwhile. Both noted that the modeling was done each year and that the results was to obtain a "significantly improved the fit to the survey biomass, especially to Arrowtooth flounder and yellowfin sole." Dr. Pilling noted that alternate models had been assessed and that the results had been presented internationally. He also suggested that "The work may benefit from developing an underlying metabolic hypothesis to drive model development. Anecdotal evidence suggests that temperature affects the movement levels of flatfish in the presence of the trawl gear. Brown *et al*. (2004) present a number of formulations linking the metabolism of organisms (linked to temperature, body mass etc.) to the ecology of populations, communities, and ecosystems. This underlying metabolic theory could be used to develop testable hypothesis models of the impact of temperature on catchability, linked to fish movement and population density."

Dr. Medley commented on the methods of fecundity estimation for the yellowfin sole stock assessment, stating that other approaches to model the size-fecundity relationship would be valuable, particularly in regards to the contribution of older females. He saw "little reason to rely on a fecundity proxy (maturity * weight) when an estimate is available." However, given that the fecundity work had already been completed he felt that it "appears reliable".

All three reviewers commented on the value of understanding productivity using the Ocean Surface Current Simulation (OSCURS) model. They noted that the model has been" validated from predictions of the landfall (time and area) of spilled cargo from container ships", and that a "1988 regime was identified for rock sole and arrowtooth from their stock-recruitment data." They found that the larval dispersal model was an interesting scientific study. Dr. Pilling made extensive comments on the difficulty in using the larval distribution predications and tying them to future stock assessment models because of the subsequent post settlement processes that could greatly change the patterns of juvenile and sub-adult abundance and distribution. He states that "improving the understanding of the relative impact of key stages of development on recruitment (e.g. Nash and Dickie-Collas, 2005) is a worthwhile pursuit.", and should also be examined under a Management Strategy Evaluation. Dr. Medley noted that, even though he saw the value of this modelling effort, that "this is likely to prove a difficult area to build reliable quantitative relationships as there will probably be a number of mutually dependent non-linear factors such as ice shelf extent, distribution of habitat, primary production as well as bottom temperature determining dynamics and availability of fish to the fishery."

Dr. Chen and Dr. Medley specifically addressed the weakness in using ecosystem-based models. Dr. Chen was concerned about the catchability log-linear model, $q = e^{\alpha+\beta T}$ with bottom temperature. Although the fit was good to the data, he stated that a "more appropriate and realistic model would be a logistic model." He felt that the form of a relationship should be justified if it is used in the model. He was also concerned that a putative "regime shift in 1988/89 is used to model the stock-recruitment for rock sole and Arrowtooth." Although there seems to be "higher productivity for the earlier years from 1977 to 1988 than the late years from 1989 to 2000", it is important to test this observation statistically.

Finally, Dr. Medley expressed his concern that sexual differences were not being handled properly for arrowtooth flounder. His concerns included the potential difference in selectivities related to length that may occur, that fishing mortalities (F_{mi}, F_{fi}) could be different, and that "the same natural mortality at unit length can produce different natural mortalities for different sexes if their growth varies". He suggested the use of the Lorenzen model even though when used "with separate sex mortalities fitted the sex ratio only slightly better than the fixed mortality model used in the assessment." However, because "fishing mortalities and selectivity may be different between the sexes" it should be "explored as part of the full stock assessment model."

TOR 3- The harvest control rules adopted for Bering Sea yellowfin sole and northern rock sole (where a stock-recruitment model and FMSY quantities are estimated) compared to other flatfish stocks where proxy values are used. Specifically, comments on the trade-offs between the different approaches are required. [Note that Rex sole appears in the reviews also].

A system of six tiers is used to choose harvest strategies for Alaskan flatfish, which include several choices in setting TACs and OFLs. A stock is assigned to a tier depending on the extent and quality of data available. Harvest strategies are more conservative for stocks that have less data or data of poorer quality. Stocks are assigned to tier 1 when there are complete statistical models, reliable point estimates of B and B_{msy}, and when there are reliable estimates of F_{msy}. In contrast, stocks are assigned to tier 6 when the only data available is catch history. The tier system was designed to allow greater levels of exploitation as data on the stocks is improved and all three reviewers noted that the tier system encourages better data collection and reporting. As noted in this report by several reviewers, "This system has multiple and cumulative layers of conservation for a high levels of biomass."

Strengths of this management system include inclusion of precautionary measures, encouragement of the industry to improve data quality, and excellence of the stock assessment for several of the species. The reviewers believed that the tier system applies "precaution in dealing with structural uncertainty, such as on the reliability of the stock recruitment relationship." They believed that the principles behind the tier system were excellent and that the harvest rules were appropriate.

The reviewers specifically commented that "Tier 1 is supported by sound research with reliable point estimate of B and Bmsy and reliable pdf of Fmsy." Based on the quality of data the all agreed that the stock assessment for BSAI yellowfin sole and rock sole showed as "good a fit as possible to the data available" and that there were reliable estimates of B, B_{MSY}, and F_{MSY}. Thus there was consensus that this species should be managed under Tier 1. For yellowfin sole, although the SSC had indicated that the comparison of stock-recruitment should be done differently between climatic periods, the reviewers believed that the stock assessment team had tested correctly.

The weakness of this management system included the lack of documentation of how decisions were made subsequent to the assessments, unnecessary complexity of the system of tiers, and assumptions surrounding M. Several of the reviewers commented on the lack of documentation and that "there is no substantive research supporting the use of the tiers or the definitions of F_{OFL} or F_{ABC} within each tier", with the exception of tier 1. Further, all three reviewers expressed concern with obtaining reliable estimates of M that would categorize flatfish species as tier 4 or 5. They mentioned that M was assumed to be 0.2 for several species and that this seemed unreliable given that when M for other flatfish species is estimated within the stock assessment models, M is lower than 0.2. Given this concern, they stated that some species now assigned to tiers 4 and 5 should be in tier 6. Finally, all three reviewers were concerned that F reference points assumed fixed selectivity and that this was very problematic for Rex sole. Currently high F is accepted because catch consist of old, large fish. However if selectivity were to change to target younger fish, then this would "lead to overexploitation at that high F."

Recommendations

I present herein a synthesis of the recommendations presented in the individual reviews. All reviewers did not address the same points, so I have indicated the number of reviewers endorsing each recommendation.

TOR 1. Modeling efforts for Bering Sea/Aleutian Islands and Gulf of Alaska flatfish assessments and harvest recommendations.

TOR 1a. The analysts' use of fishery dependent and fishery independent data sources in the assessments

Recommendation 1. Examine the commercial and observer time series of catches as an alternative indicator of abundance for use in the future when survey time series may become limited. In following this recommendation, it is important to 1) undertake "standardisation, using covariates such as trawl net type, depth, temperature, tow length, vessel speed and so on"; 2) "generate observer-based abundance indices for the flatfish to run alongside the current survey index." (PM, GP)

Recommendation 2. The studies on survey catchability have been successful and should be continued. As well as improving the survey indices, this research will prove invaluable as the basis for standardizing the observer data. (PM)

Recommendation 3. Determine the robustness of the outputs of flatfish assessment approaches to uncertainty in commercial and observer catch data, using Management Strategy Evaluation. For example, identify the level of data uncertainty at which assessments become biased or imprecise, and whether this level is felt representative of the current data available. If so, how can the data collection be improved (e.g. how complete does observer coverage need to be)? (GP)

Recommendation 4. Current assumptions of minimum biomass outside US waters (e.g. Greenland turbot) should be evaluated, and investigations pursued where existing knowledge suggests the level of catches in neighbouring waters may be important. In particular, AFSC scientists should improve their communications and co-operation with Russian scientists, if possible, sharing survey and catch information. (PM, GP)

Recommendation 5. Management Strategy Evaluation (MSE) should be used to 1) examine the importance of collaboration on catch and abundance data between U.S. and Russian scientists, 2) "to examine the significance of different migration hypotheses on the effectiveness of management."; 3) develop a process-based ecosystem model. with the intention of conducting more extensive management strategy evaluations. "The MSE will need to consider how to measure performance of the management strategies being tested, through considering utility and regret functions, ecological and multispecies indicators as well as more traditional fisheries indicators and reference points." (PM, GP)

Tor 1b. Gaps or inconsistencies in the population dynamics modeling methodology or logic;

Recommendation 6. Routine performance of retrospective analyses of the stock assessments is recommended, which are included in the stock assessment reports. (PM, GP)

Recommendation 7. Given the influence of the form of stock-recruitment relationship (particularly for rock sole and arrowtooth sole) used within assessments, scientists should 1) perform sensitivity runs or preferably Management Strategy Evaluation to look at the robustness of current assessments and management to this uncertainty, 2) use a "rigid statistical approach" to test whether the temperature shift is real (see Dr. Chen report for details), 3) examine model temperature effects on recruitment; 4) incorporate temperature directly into the SR model (see Dr. Chen report for details). (DC, GP)

Recommendation 8. Age-specific natural mortality estimates should be developed using the MSVPA estimates of predation mortality at age, or Lorenzen's formula, to develop biologically more realistic natural mortality-at-age estimates, along with informative priors. Examine the impact of this on the assessment results, stock recruitment relationship and current tier 4/5 calculations. (PM, GP)

Recommendation 9. Examine the implications of using separate models for males and females for mean weight-at-age within current assessments, potentially using Management Strategy Evaluation. (PM, GP)

Recommendation 10. Include estimates of density dependent growth if it is present, and more accurate estimates of population parameters such as fecundity. (PM)

Recommendation 11. Continue to develop multispecies projection models (particularly with regard to flatfish as halibut bycatch) to allow more explicit consideration of the interactions between species and fisheries in management advice. "If sex-disaggregated assessment models are found to be beneficial, sex-specific projection models should be considered." (DC, GP)

Recommendation 12. The impact of the assumptions made and settings selected when fitting the SS2 model to the time periods of catch and survey data available for Greenland turbot should be examined. If this has already been performed, it would be useful to include a short summary of these within the assessment report for future reviews. (GP)

Recommendation 13. The 2005 analysis of Gulf of Alaska rex sole was based on intermittent age information up to 1996. Given notable changes in length-at-age have been seen in historical data, available age information for years post 1996 should be processed to ensure any trends in length-at-age are incorporated within the assessment. (GP)

Recommendation 14. For Bering Sea and Aleutian Islands yellowfin sole, some uncertainty exists over the accuracy of early ages based upon whole otolith readings. Examine the impact of relaxing model fitting constraints on early years of the catch age structure time series for yellowfin sole on stock assessment results. (GP)

Tor 1c. If uncertainties in assessment model results are appropriately applied to management advice

Recommendation 15. Even though the assessment and projection conform to National Standard Guidelines, provide a better indication of uncertainty within the assessment and projection model, as for example inclusion of uncertainty in current year numbers-at-age and use of the estimated stock-recruitment relationship with accompanying uncertainty. (GP)

Recommendation 16. Sensitivity analyses on alternative initial parameters and local optimization should be investigated with ADMB. (DC)

Recommendation 17. More sensitivity runs are needed for the flatfish assessments to look for the assumptions which really do make a difference (e.g., structural, statistical, assumed fixed parameters, priors used). (DC)

Recommendation 18. Include parameter uncertainty which is generated from the ADMB in the projections to generate the projected biomass distribution instead of a point estimate. (DC)

Recommendation 19. Assess potential for parameter confounding between the steepness and other parameters. (See Dr. Chen section 4.3 which is rich in detail) (DC)

TOR 2. The effort to incorporate ecosystem indicators and shifts in states of nature in the assessments.

These include modeling survey catchability with annual bottom water temperature and using the Ocean Surface Current Simulation Models (OSCURS) to define putative oceanic productivity regimes.

Recommendation 20. The entire CIE panel recommended that more research and effort should be made in the development of multispecies/ecosystem modeling and management. (DC, PM, GP)

Recommendation 21. Examine the use of metabolic theory to develop testable hypotheses to better explain the impact of temperature on survey catchability. (GP)

Recommendation 22. Develop a detailed process ecosystem model that could 1) examine the impact of temperatures on fleet distributions using available VMS data. 2) include multiple representations of the main ecological groups and processes found in marine ecosystems, 3) include a sampling model that reports fisheries statistics and the results of fisheries independent data collection exercises, with error structures typical of that found in reality, 4) model fleet-level to individual-level representation of the dynamics of human sectors (primarily fisheries), 5) include assessment models (from simple estimates to the explicit assessment models used in reality; biomass dynamics models, catch-at age models etc) 6) include management decision models, including lag in implementation of decisions; and simple to complex representations of the socio-economic drivers and considerations (particularly for compliance) and 7) apply a management strategy evaluation (MSE) to its results. (PM, GP)

Recommendation 23. “Develop research programmes to identify and understand the key developmental stages for flatfish and their associated processes, and/or identify those stages key to assessment and management through Management Strategy Evaluation in order to focus this research. This should aim to link the larval dispersal model results to recruitment levels.” (GP)

Recommendation 24. “Examine routes to progress multispecies advice from the qualitative advice currently given (which it must be noted is already a step beyond many fisheries organisations!), to more integrated quantitative advice and multispecies projection models tie in directly with current assessment methods and prescribed management approaches, while studies provide a wider view of the multispecies ecosystem in the region that may be appropriate for future management requirements.” Both MSVPA and Ecopath with Ecosim approaches should be continued. (GP)

Recommendation 25. Develop a recruitment index measured subsequent to density-dependent mortality that can be used to improve projections and short term forecasts. (PM)

Recommendation 26. Monitor juveniles and adults before maturity to allow consideration of the effects of climate on growth, growth rate and population density, and use these data in future models. (PM)

TOR 3. The harvest control rules adopted for Bering Sea yellowfin sole and northern rock sole (where a stock-recruitment model and FMSY quantities are estimated) compared to other flatfish stocks where proxy values are used. Specifically, comments on the trade-offs between the different approaches are required. [Note that Rex sole appears in the reviews also].

Recommendation 27. Simplify the Harvest Control Rule 6-tier system for flatfish. (Consensus DC, PM, GP)

Recommendation 28. Use Management Strategy Evaluation to ensure that the tier decision system is working as expected with respect to the robustness and precautionary nature of advice in given data situations, and to improve the method to recommend the OFL and ABC. (PM, GP)

Recommendation 29. “Examine the performance of the harmonic mean estimate of F_{msy} under current and alternative assumptions of model uncertainty.” (GP)

Recommendation 30. Examine the implications of uncertainty in natural mortality on the performance of tier 5 management rules and incorporate uncertainty in the HR strategy tier system and bring the loss function into the tier system. (GP)

Recommendation 31. Classify yellowfin sole and northern rock sole as tier 1 based on the reliable estimates from the assessment model. (DC)

Recommendation 32. Based on the collection of new age information, analyses of alternative contributing factors (e.g. the use of excluders in gear) etc., examine the implications of alternative selectivity and maturity ogives on assessment results for Gulf of Alaska rex sole. (GP)

III. Appendices

References

Goodman, D., M. Mangel, G. Parkes, T. Quinn, V. Restrepo, T. Smith, K. Stokes (With assistance from G. Thompson). 2002. Draft Report. Prepared for: North Pacific Fishery Management Council, November 21, 2002.

Ianelli, J.N., T.K. Wilderbuer, and D. Nichol Assessment of Greenland Turbot in the Eastern Bering Sea and Aleutian Islands

Turnock, B.J., T.K. Wilderbuer and E.S. Brown. Gulf of Alaska Arrowtooth Flounder Stock Assessment

Turnock, , B.J., T.K. Wilderbuer and E.S. Brown. Gulf of Alaska Flatfish

Turnock, B.J., and Z.T. A'mar. NMFS Alaska Fisheries Science Center Gulf of Alaska Rex Sole Stock Assessment

Wilderbuer, T.K. and D.G. Nichol. BSAI Northern Rock Sole

Wilderbuer, T.K. and D.G. Nichol. BSAI Yellowfin Sole

Wilderbuer, T.K., D.G. Nichol and P.D. Spencer. Other Flatfish

Statement of Work

Consulting Agreement between Dr. Cynthia Jones and NTVI

Alaska Flatfish Review Statement of Work

General

The Alaska Fisheries Science Center (AFSC) requests a peer review of the Gulf of Alaska and Bering Sea/Aleutian Islands flatfish stock assessments by the Center for Independent Experts (CIE). Flatfish fisheries in Alaska are valuable and are likely to receive increased interest from commercial fishers in the next decade. Most flatfish populations are currently at stable or high levels of abundance. The North Pacific Fisheries Management Council (Council) is likely to pass amendments to rationalize (establish quota shares) groundfish fisheries in the near future. These amendments will allow for more flexibility in the time and areas fished which may result in the development of renewed interest in flatfish fisheries. Thus the AFSC desires an independent peer review of these stocks to assess the quality of the assessments and to determine whether the Council is being provided with the best available information and analysis.

There are currently 21 flatfish species managed under the Bering Sea and Aleutian Islands Fishery Management Plan, of which six species have age-structured stock assessments. In the Gulf of Alaska there are ten managed species, of which four have age-structured assessments. These assessments have never undergone a CIE review. While the current Council review process is in place, there remains a compelling need for an independent peer review of the Alaska flatfish assessments.

The CIE assessment review requires a total of four reviewers who are thoroughly familiar with various subject areas involved in stock assessment, including population dynamics, separable age-structured models, harvest strategies, survey methodology, and the AD Model Builder programming language. They should also have experience conducting stock assessments for fisheries management. The reviewers will travel to Seattle, Washington, to discuss the stock assessments with the flatfish assessment authors and other scientists at the AFSC involved in flatfish stock assessment. Overview presentations will be made on the topics listed below and assessment authors will be available for questions from reviewers. Specific details on a few (3-4) selected stock assessments will be presented to assist in the review.

Three of the reviewers shall generate individual reports. The fourth reviewer shall generate a Summary Report that compiles the points made by the three individual reviewers into one succinct document. The individual reports shall be appended to the Summary Report, thereby providing the complete detailed information from the individual reviewers. All reports shall address the strengths and weaknesses of the following points.

- Modeling efforts for Bering Sea/Aleutian Islands and Gulf of Alaska flatfish assessments and harvest recommendations. Specifically, the review shall evaluate:

    - The analysts' use of fishery dependent and fishery independent data sources in the assessments;
    - Gaps or inconsistencies in the population dynamics modeling methodology or logic;
    - If uncertainties in assessment model results are appropriately applied to management advice; and
    - Whether the assessments provide the best available science.

- The effort to incorporate ecosystem indicators and shifts in states of nature in the assessments. These include modeling survey catchability with annual bottom water temperature and using the Ocean Surface Current Simulation Models (OSCURS) to define putative oceanic productivity regimes.

- The harvest control rules adopted for Bering Sea yellowfin sole and northern rock sole (where a stock-recruitment model and F_{MSY} quantities are estimated) compared to other flatfish stocks where proxy values are used. Specifically, comments on the trade-offs between the different approaches are required.

The AFSC will provide copies of stock assessment documents, survey reports, and other pertinent literature on a web site.

Specific

The CIE shall provide four reviewers, for a maximum total of 52 work days. The three individual reviewers are approved for a maximum of 15 work days each, and the summarizer is approved for a maximum of seven work days. The three individual reviewers shall be approved for travel to the meetings at the AFSC. The summarizer shall work from his or her home office, so no travel is required.

Specific requirements for the individual reviewers and the summarizer are listed below[1]. The list is followed by a chronological table, which includes the required steps for the CIE. If any intermediate step in the review process is delayed, all subsequent steps that depend on it will be delayed by an equivalent period.

Individual Reviewers

1. Read and become familiar with the relevant documents provided to the reviewers. The ten age-structured assessments that encompass the focus of the review are presented in 512 pages (of which approximately 1/3 is text).
2. Discuss the stock assessment with the lead assessment scientist and survey scientists at the AFSC, in Seattle, Washington, from June 11 to June 14, 2007 (see attached agenda).
3. No later than June 29, 2007, submit a written report of findings, analysis, and conclusions. More details on the report outline and organization are provided in

[1] All reports will undergo an internal CIE review before they are considered final.

Annex I. The report shall be sent via e-mail to Dr. David Die (ddie@rsmas.miami.edu), and to Mr. Manoj Shivlani (mshivlani@rsmas.miami.edu), and to the CIE summarizer (contact information to be provided by the CIE).

Summarizer

1. Become familiar with the documents provided to the individual reviewers.

2. Read the draft reports from the individual CIE reviewers, and draft the summary report following the standard outline. More details on the report outline and organization are provided in Annex I.

3. Complete summary report within one week of receiving final individual reviewer reports from the CIE. The report shall be sent via e-mail to Dr. David Die (ddie@rsmas.miami.edu), and to Mr. Manoj Shivlani (mshivlani@rsmas.miami.edu).

Activity	**Deadline**
Individual reviewers submit their reports to the CIE and the summarizer.	June 29, 2007
CIE reviews and approves the individual reviewer reports, and provides them to the NMFS COTR.	July 13, 2007
NMFS COTR approves individual reviewer reports.	July 17, 2007
CIE provides final individual reviewer reports to NMFS COTR and to the summarizer.	July 20, 2007
Summarizer provides summary report to CIE.	July 27, 2007

CIE reviews and approves the summary report, and provides it to the NMFS COTR.	August 10, 2007
NMFS COTR approves summary report.	August 14, 2007
CIE provides final summary report to NMFS COTR.	August 17, 2007

Submission and Acceptance of CIE Reports

The CIE shall provide the final reports for review for compliance with this Statement of Work and approval by NOAA Fisheries to the COTR, Dr. Stephen K. Brown (Stephen.K.Brown@noaa.gov), according to the above schedule. The COTR shall notify the CIE via e-mail regarding acceptance of each report. Following the COTR's approval, the CIE shall provide a pdf format version of each approved report to the COTR.

CIE Flatfish Assessment Review

NMFS Alaska Fisheries Science Center
7600 Sand Point Way NE, Building 4
Seattle, Washington

Agenda May 11 Draft version June 11-14, 2007

Monday June 11th

9:00	Welcome and Introductions, adopt agenda	
9:15	Overview (species, surveys, fishery, catch levels, ABCs, TACs, bycatch)	**Tom**
10:00	Biology (growth, natural mortality, diets, spawning & nursery areas, maturity)	**Buck, Dan, Janet**
11:00	Trawl experiments on herding, escapement, bottom contact issues	**Peter Munro**
11:30	Age Determination of flatfish	**Delsa and Craig**
12:00	***Lunch***	
13:00	Observer Program	**Jennifer Ferdinand**
14:00	Harvest control rules and projection model	**Jim**
15:00	Summary of on-going research	**Tom, Dan, Janet, Kerim**

Tuesday June 12th

9:00	**Gulf of Alaska and Bering Sea stock assessments**	
	Overview of models	**Jim**
	BSAI Yellowfin sole/northern rock sole	**Tom**
	GOA arrowtooth flounder	**Jack**
	BSAI Greenland turbot	**Jim**
	GOA rex sole	**Jack or Buck**
	GOA shallow-water flatfish	**Buck**
12:00	*Lunch*	
13:00	**Incorporating ecosystem indicators in stock assessments**	
	Modeling catchability with bottom temperature	**Tom, Paul**
	Catchability and distribution changes with temperature	**Paul**
	OSCURS model to define productivity regimes	**Tom**
	IBM 3-D drift model	**Buck**

Wednesday June 13th

9:00	Harvest control rules using F_{msy} for yellowfin sole and northern rock sole compared to proxy values	**Tom**
12:00	*Lunch*	
13:00	Ecosystem aspects of flatfish	**Kerim**

Thursday June 14th

Reviewer discussions with assessment authors

Annex 1: Contents of Reviewer Reports

The following requirements refer to all reports, both the individual reports and the Summary Report.

1. All reports shall be prefaced with an executive summary of findings and/or recommendations.

2. The main body of all reports shall consist of a background, description of review activities, summary of findings, conclusions/recommendations, and references.

3. The reports shall also include as separate appendices the bibliography of all materials provided and any additional papers cited, along with a copy of the statement of work.

Summary report only: The summary report shall include the three individual reviewer reports as appendices.

Review from Dr. Din Chen

CIE Review Report for Alaska Flatfish Assessment Review

Dr. Din Chen

Fisheries Statistics Consultant

Prepared for

Center for Independent Experts

EXECUTIVE SUMMARY

The Alaska Fisheries Science Center (AFSC) requested a peer review of the Gulf of Alaska (GOA) and Bering Sea/Aleutian Islands (BSAI) flatfish stock assessments by the Center for Independent Experts (CIE). From June 11 to 14, 2007, a CIE panel consisting of Drs. Din Chen, Paul Medley and Graham Pilling, met at the AFSC to review the flatfish stock assessments and other related documents prepared by a group of AFSC scientists (Appendix 1: List of Participants).

There are currently 21 flatfish species managed under the Bering Sea and Aleutian Islands (BSAI) Fishery Management Plan, of which six species have age-structured stock assessments. In the Gulf of Alaska (GOA) there are ten managed species, of which four have age-structured assessments. These assessments have never undergone a CIE review.

The review meeting started on Monday, June 11 (Appendix 2: Agenda) with the overview (for the species, surveys, fishery, catch levels, ABCs, TACs and bycatch), fish biology (on fish growth, natural mortality, diets, spawning, nursery areas, maturity), ageing (age determination and validation of flatfish), survey efficiency (trawl experiments on herding, escapement, bottom contact issues), the Observer program, summary of ongoing research and the harvest control rules and projection models.

The major species reviewed on Tuesday June 12 were BSAI yellowfin sole/northern rock sole, GOA arrowtooth flounder, BSAI Greenland turbot, GOA rex sole and GOA shallow-water flatfish (Please refer to the agenda in Appendix 2). Model review was started from an overview by Dr. Jim Ianelli on the age-structure assessment models for these flatfish species. Issues on incorporating ecosystem indicators in stock assessments were discussed by the CIE panel and the assessment team in the afternoon on modeling catchability with bottom temperature, catchability changes with temperature and distribution changes with temperature.

On Wednesday June 13, the CIE panel and assessment team discussed the Ocean Surface Current Simulation (OSCURS) model on defining productivity regimes, the IBM 3-D drift model and the harvest control rules using F_{msy} for yellowfin sole and northern rock sole compared to proxy values. On Thursday morning of June 14, the CIE reviewers discussed with the assessment authors to clarify some issues from the assessment meeting and adjourned the meeting at noon.

Based on the review and discussion of the stock assessment document and related presentations, the modeling efforts for BSAI and GOA flatfish assessments and harvest recommendations were generally appropriate in using the fishery dependent and fishery independent data sources in the assessments. The age-structure assessment model with ADMB was appropriately applied for flatfish population dynamics except the using of SS2 with steepness parameter re-parameterized from Ricker and B-H stock-recruitment model to the Mace-Doonan formulation, which has been known to be confounded with other parameters (see Recommendations). It is concluded that the reviewed assessments provide the best available science with the harvest control rules based on the tier system.

Furthermore, it is considered a useful and valuable effort to incorporate ecosystem indicators and shifts in states of nature in the assessments, which included modeling survey catchability with annual bottom water temperature and using the Ocean Surface Current Simulation Models (OSCURS) to define putative oceanic productivity regimes.

In summary, the scope of this CIE review was quite broad covering the entire modeling efforts for BSAI and GOA flatfish assessments for the input data, supporting science, analytical methods, projections and harvest strategies and harvest recommendations. Thousands of pages of background information were provided on two FTP sites.

Even though these stocks are not commercially important at present, a great deal of effort by the AFSC has been put into these stock assessments. The AFSC assessment team is commended for the quality of the documents and presentations provided for review and their cooperation requested during the meeting. The flatfish team did an excellent job of presenting a wide-array of relevant information including many aspects of the current research programs, stock assessment models and harvest strategies.

1. BACKGROUND

Designated by the Center for Independent Experts (CIE), the author was invited as a panelist (Appendix 4: Statement of Work) to the Alaska Flatfish Review to review the stock assessments for Alaska flatfish.

On June 1, 2007, Mr. Manoj Shivlani, the CIE Coordinator, provided the author and other panel members with a website (http://www.afsc.noaa.gov/REFM/Stocks/assessments.htm) to the background information of the stock assessment documents and the associated documents for this review. An extensive list with thousands of pages of documents was listed on this ftp site for all the species from BSAI and GOA. On June 4, 2007, a succinct list of review documents with the early version of meeting agenda were provided by Tom Wilderbuer from AFSC with an ftp site (ftp://ftp.afsc.noaa.gov/afsc/public/flatfish/flats.html). Although there was limited time, I read the review documents and consulted the background materials from the provided ftp sites and some related AFSC web sites.

I would like to comment that Mr. Manoj Shivlani of the CIE did an excellent job arranging the contract, hotel and flight, and making everything run smoothly.

2. STOCK ASSESSMENT REVIEW ACTIVITIES AND PANEL DISCUSSIONS

The meeting to review the Alaska flatfish assessments took place in AFSC, Seattle, Washington, from June 11 to 14, 2007. During the meeting, an ftp site was made available to facilitate the sharing of presentations and other related documents.

The meeting convened at 9:00am on June 11. Patricia Livingston, the Director of Resource Ecology and Fisheries Management Division, welcomed the group. Tom Wilderbuer, the review meeting coordinator, then opened the meeting with an overview of the flatfish fisheries and management which included the species with the management categories, the fishing areas, catches, relative stock sizes and the council/assessment process with the management tier system. Thereafter, a series of well-organized presentations were given from the Assessment Team and panel discussions followed from each presentation. Appendix 3 describes the detailed review activities and panel discussions.

3. SUMMARY OF FINDINGS

This part is organized according to the terms of reference (TOR) in the review statement of work (SOW) (Please refer to Appendix 4 for the SOW) to address the strengths and weaknesses of the following TORs. The bullet points are the key findings, followed by a detailed discussion, if any.

TOR 1: Modeling efforts for Bering Sea/Aleutian Islands and Gulf of Alaska flatfish assessments and harvest recommendations. Specifically, the review shall evaluate:

- **The analysts' use of fishery dependent and fishery independent data sources in the assessments;**
- **Gaps or inconsistencies in the population dynamics modeling methodology or logic;**
- **If uncertainties in assessment model results are appropriately applied to management advice; and**
- **Whether the assessments provide the best available science.**

Generally, the modeling efforts for BSAI and GOA flatfish assessments and harvest recommendations are appropriate given the available data. The assumed age-structure population dynamics are consistent with the current understanding and knowledge of the flatfish biology. The appropriateness of assessment input data and the analytical approaches have been reviewed extensively in previous workshops and reports. The quality of the harvest recommendations rely on good data and methods and the modeling uncertainties with Bayesian approach and simulations are mostly appropriate. The assessments of these flatfish species generally provide the best available science for management advice.

Strengths

- Ageing data are well determined and validated with several methods. The ageing methods are well respected and produced very reliable data.
- Long catch history time series.
- Extensive trawl survey data available to generate fishery independent data.
- Wider coverage for the Observer program.
- The likelihood based estimation with ADMB is adequate for the age-structure population dynamics

Weaknesses

- The trawl surveys have undergone several changes in standardization and operation. They were conducted annually on the Bering Sea shelf, semi-annually in the Gulf of Alaska and the Aleutian Islands. Budget constraints have reduced trawl surveys on the Bering Sea slope.
- Natural mortality (M) estimate is a notoriously difficult parameter to estimate. For the flatfish stocks, this M is even problematic and used 0.2 for most of the stock. This is especially problematic for the stock in tier 5.
- Migration and distribution patterns among EBS, WBS and GOA are generally assumed not to exist, but limit tagging studies did show mutual migration.
- There are a larger number of parameters estimated from the assessment models; parameter confounding and local optimization are bound to be a problem. More investigation is required to address this issue. For example, there is a known confounding effect for the steepness parameter and other parameters in SS2 assessment model.

Fishery independent data

Trawl surveys represent the key fishery independent data for the flatfish assessment. These are used to obtained at-age and at-length data for the assessment. There are extensive trawl survey data from BSAI and GOA. The surveys were conducted annually on the Bering Sea shelf and semi-annually in the Gulf of Alaska and the Aleutian Islands. It is unfortunate that budget constraints have reduced trawl surveys on the Bering Sea slope. To ensure continuous and valid assessments for the future, NMFS should ensure the continuation of these surveys and their operations.

Fishery dependent data

There exist long time series data from fishery catches and the associated Observer sampling program appears well-founded and executed. Observer data provided validation of catch composition, bycatch, effort, location, and obtained biological metrics and collections. The sampling methods and coverage are appropriate and well documented. Considerable effort has gone into training and quality control. The Observer program had coverage from 200% to 0%. There was 200% coverage for vessels 125 ft and greater in length that was fishing Pollock, with two observers on board and all fishing events sampled. There was 100% coverage for vessels 125 ft and greater in length, with one observer on board and random selection of fishing events sampled. There was 30% coverage (for 20% of the vessels) for vessels between 60 and 125 ft in length, with 30% of fishing days sampled by random selection of fishing events. There was 0% coverage (for 72% of the vessels) for vessels under 60ft in length. The priorities for the Observer program are to record fishing effort information for every haul (set) retrieved, sample for official total catch estimates and species composition, and send data to the observer program in Seattle, Washington.

Migration issues

The flatfish stocks are generally assumed to have little migration in BSAI and GOA. The assessment models were carried out without any migration component. The CIE panel noticed and questioned this issue, especially between EBS and Russian waters.

There are several ad hoc tagging studies which showed migration among these areas, which might indicate that the assumed population dynamics are too simplified. Any violation of this assumption can lead to misinterpretation of abundance data and unreliable stock assessments.

Assessment models

The assessment models for these flatfish species are based on the standard age-structured population dynamics models, which are considered appropriate without detailed information on stock structure and migrations. When more information is available, it may be more appropriate to incorporate the migration component into the existing assessment models.

Stocks, such as yellowfin sole, etc, should be considered using two-sex models since there are large growth differences between the sexes. The current practice is to assume that there are sexually explicit differences in growth and weight at age.

For parameter estimation, the current estimation methods can be described as “quasi-Bayesian”. The estimates are derived by minimizing a negative log-likelihood modified by some prior distributions. From my experience in Bayesian modeling (I actually did my Ph.D. in Bayesian statistics), it should be prudent for any Bayesian modeling for sensitivity investigation for the appropriate choices of the prior distributions and the associated parameters (Millar, 2004).

There is a larger number of parameters estimated from the assessment models; thus, parameter confounding and local optimization are bound to be a problem. More investigation is required. For example, there is a known confounding effect for the steepness parameter and other parameters in SS2 assessment model (see 'Recommendations').

Modeling of uncertainty and application to management advice

The principle of any assessment is to model the fish population dynamics with available biological data and to capture the appropriate level of uncertainty from the model process and data measurement. This is done to ensure that assessment results are more realistic in terms of the "true" uncertainty but are still useful for management purposes.

In general, the flatfish assessments are appropriate to capture the proper level of uncertainty with the ADMB MCMC and bootstrapping.

TOR 2: The effort to incorporate ecosystem indicators and shifts in states of nature in the assessments. These include modeling survey catchability with annual bottom water temperature and using the Ocean Surface Current Simulation Models (OSCURS) to define putative oceanic productivity regimes.

The ecosystem based approaches to incorporate ecosystem indicators and environmental regime shifts into flatfish stock assessment is intriguing in modeling survey catchability with annual bottom water temperature and using the Ocean Surface Current Simulation Models (OSCURS) to define putative oceanic productivity regimes.

<u>Strengths</u>

- In modeling catchability with bottom, the survey catchability (*q)* is modeled each year by considering the relationship to annual bottom water temperature which significantly improved the fit to the survey biomass, especially for arrowtooth flounder and yellowfin sole.
- In modeling flatfish distribution changes with temperature, a strong statistical relationship between fish distribution and cold pool distribution have been found for several flatfish species.
- Productivity regimes were defined using the Ocean Surface Current Simulation (OSCURS) model. The essence of the OSCURS model is to convert the selected daily sea level pressure grid to east-west wind velocity components and then current grids using empirical functions. The model has been validated from predictions of the landfall (time and area) of spilled cargo from container ships. A 1988 regime was identified for rock sole and arrowtooth from their stock-recruitment data.

Weaknesses

- The log-linear model, $q = e^{\alpha + \beta T}$, to the catchability (q) with bottom temperature (T) increased the goodness of fit to the survey biomass in the observed data range. A more appropriate and realistic model would be a logistic model. The assessment authors had investigated this model and further investigation should be carried out with other relationships for all flatfish species. This evaluation may assist in determining whether the effect is due to changes in activity or changes in distribution. The form of the relationship and how it is incorporated into the model should be justified.
- A potential regime shift in 1988/89 is used to model the stock-recruitment for rock sole and arrowtooth. It was found that there existed higher productivity for the earlier years, from 1977 to 1988, than the late years, from 1989 to 2000. A rigid statistical test would be necessary to test whether this is real rather than a result of observational errors (see Recommendations for stock recruitment modeling).

General

Multispecies/ecosystem modeling and ecosystem management is one of the high priorities in fisheries research and management. It is obvious that the ecosystem management can improve the system for non-target species management, and it can incorporate ecosystem considerations into individual stock assessments to define ecosystem level reference points. This system can incorporate predator/prey data and multispecies and ecosystem models with improved links to bottom-up processes.

There are strengths and weaknesses for the ecosystem-based approaches to incorporate ecosystem indicators and environmental regime shifts into flatfish stock assessment, as previously listed.

In using the ecosystem-based approaches, the trophic approaches are more appropriate for data-rich situations, whereas in data-poor situations such as with Alaska flatfish, environmental indices may provide the best insights. The CIE panel recommended that more research and effort should be made in the development of multispecies/ecosystem modeling and management.

TOR 3: The harvest control rules adopted for Bering Sea yellowfin sole and northern rock sole (where a stock-recruitment model and F_{MSY} quantities are estimated) compared to other flatfish stocks where proxy values are used. Specifically, comments on the trade-offs between the different approaches are required.

The tier system is used for the current harvest strategies for Alaskan flatfish, and comprehensive and several subjective choices are involved in setting TACs and OFLs. There are several, specific strengths and *weaknesses in this management system for Alaska flatfish, including Bering Sea yellowfin sole, northern rock sole and other flatfish stocks.*

Strengths:

- There are six tiers for the management system depending on information available, which ranges from complete statistical models of the stock and reference points (Tier 1) with reliable point estimates of B and Bmsy and reliable pdf of Fmsy, down to stocks for which there are essentially no data (Tier 6) with just catch history. This system is used to define the ABC and OFL for the Alaska flatfish stocks, and it has multiple and cumulative layers of conservation for high levels of biomass.
- Tier 1 is supported by sound research with reliable point estimates of B and Bmsy and reliable pdf of Fmsy.
- The tier system stimulates and encourages the fishing industry and researchers to collect more data in order to move up the system with more information
- It appeared that tier 1 management is reasonable for BSAI yellowfin sole as there exist reliable point estimates of biomass and BMSY and a reliable pdf of FMSY. Appropriate stock-recruit data were used to establish FMSY and BMSY from the period 1978-2000 (i.e., post 1977 regime shift). A Ricker SR fit appeared quite reasonable and FMSY is well estimated for the stock, which is at a very high and stable level and capable of sustaining higher harvests than have been taken in the past decade.
- Similarly, it appeared that the northern rock sole could be classified into tier 1 since there are reliable point estimates of biomass and BMSY and a reliable

pdf of FMSY. This stock is at a high, stable level, has been relatively lightly exploited recently, post 1978 data show a strong Ricker-type stock-recruit relationship, and FMSY estimates are credible.

Weaknesses:

- Tiers 4-5 require a reliable point estimate of B. Several species of flatfish classified as tier 5 which is assumed to have reliable estimates of natural mortality rate M. However, this M is usually assumed to be 0.2 which in fact is very problematic and questionable.
- The CIE panel recommends simplifying the system by perhaps basing it just on the availability of reliable abundance indices.

General

The 6-tier system is used to define the ABC and OFL of Alaska flatfish stocks. It involves multiple-layers of conservation. In my reading, the assessment authors firstly recommend an ABC after several assessment runs from the author selected assessment models. Then the Plan Team recommends an ABC, which can differ from the assessment author's recommendation. Next, the SSC makes an ABC recommendation (another subjective choice), and finally the Council accepts one of the ABC recommendations and then sets a TAC at a level up to the ABC. For in-season management, the fisheries managers will then try to manage the fishery to the TAC and will certainly try to avoid any catch in excess of the ABC. There will be no directed fisheries on a stock after its TAC has been exceeded. Fisheries on one stock can be closed if the bycatch on another stock would cause the TAC of the bycatch stock to be exceeded.

It is recommended that the 6-tier system for Alaska flatfish stocks be simplified. The current structure has six levels based on different levels of available information. However, apart from tier 1 (Thompson 1999), there is no substantive research supporting the use of the tiers or the definitions of F_{OFL} or F_{ABC} within each tier. I believe that the system has been and will be successful in conserving fish stocks.

Yellowfin sole can be classified as tier 1 based on the reliable estimates from the assessment model. Tiers 4 and 5 are based on a reliable estimate of B and tier 5 is based on a reliable estimate for M. In these cases, I think that the “reliable” estimate of B must be coming from a trawl survey or other type of survey, which again is problematic since a trawl survey does not provide reliable estimates of biomass. The M estimate is more problematic. And therefore, the tier 4 and 5 might be just tier 6.

4. RECOMMENDATIONS

4.1. General recommendations

a. Natural mortality estimates should be reviewed. This is especially important for tier 4/5 stocks. Informative priors could be developed along with this line. Dr. Paul Medley in his report will comment on using MSVPA and Lorenzen model to estimate M.
b. Sensitivity analyses on alternative initial parameters and local optimization should be investigated with ADMB. My experience for ADMB is that it is an excellent software for fishery stock assessments, but it is essentially software for local optimization where different initial parameters could give very

different convergences. I have run into this local optimization very often with a much lower number of parameters than the model used for this assessment. I have raised this point in the Review meeting and the Assessment Team will investigate the parameter estimates with different starting values.

c. More sensitivity runs are needed for the flatfish assessments to look for the assumptions which really do make a difference (e.g., structural, statistical, assumed fixed parameters, and priors used).

d. Parameter uncertainty which is generated from the ADMB in the projections to generate the projected biomass distribution instead of a point estimate should be included.

e. Uncertainty in the HR strategy tier system should be incorporated, as well as bringing in the loss function into the tier system. Dr. Paul Medley will comment on this in detail in his report.

4.2. Stock-recruitment model incorporating regime shift or temperature, etc.

In stock-recruitment (SR) models for the flatfish species such as rock sole and arrowtooth, incorporating regime shift or bottom temperature allows the SR model to fit within the assessment model. I do not have full access to the assessment models (I did get the SR data from Dr. Tom Wilderbuer and could fit the SR model outside the assessment model, but that would be slightly different) and therefore I will provide some conceptual models to be included into the future assessment models as follows:

a. Regime SR model (models 13 to 15 in Chen 2001)

A potential regime shift in 1988/89 was used to model the stock-recruitment for rock sole and arrowtooth. It was found that there existed higher productivity for the earlier years, from 1977 to 1988, than for the late years, from 1989 to 2000. A rigid statistical approach to test whether this shift is real and has a significant effect to the SR models can be formulated by using a dummy variable defined as T = 1 for years from 1977 to 1988 and otherwise 0 for years from 1989 to 2000. With this dummy variable, the two regime SR models can be cast into one formulation as:

$$R = S \times \exp(\alpha + \beta S + \alpha_0 T + \beta_0 ST + \varepsilon):$$

If the regime shift is shown to affect the SR relationship, then α_0, β_0 should be both statistically significantly different from zero.

If α_0 is statistically significantly, but β_0 is not, then there exists a significant effect for the productivity parameter, but not to the density-dependent parameter.

If α_0 is not, but β_0 is statistically significantly, then there exists a significant effect for the density-dependent parameter, but not on the productivity parameter.

If both α_0 and β_0 are not statistically significant, then the regime shift does not affect the SR relationship.

b. Incorporating the temperature directly into the SR model

CIE panel raised this point in the meeting and recommended that the assessment authors incorporate the temperature data directly into the SR model as:

$$R = S \times \exp(\alpha + \beta S + \gamma Temp + \varepsilon)$$

This is exactly model 12 in Chen (2001). A slight extension to this direction is model 16. In this paper Chen (2001) demonstrated the fuzzy logic SR model is preferred over all other models.

If the assessment team is interested in this approach, I will gladly provide guidance incorporating the fuzzy logic SR model into the flatfish assessment.

4.3. Parameter confounding

In the modeling exercise, the flatfish species were mainly assessed by coding the age-structured fish population dynamics model into ADMB, and there is a tendency to move to SS2. For example, BSAI Greenland turbot was assessed using SS2 for this review. It is becoming known that there is a parameter confounding between the steepness and other parameters. The re-parameterization of the original Ricker or Beverton-Holt stock-recruitment models to the Mace-Doonan formulation in SS2 (Methot 2005, page 8, equation 1.6) may lead to numerical instabilities.

Steepness (h) and S_0 are more highly confounded in the Mace-Doonan formulation than are the α and β parameters in the original Ricker or Beverton-Holt formulation. I have raised this point in a previous (STAR panel) review and would like to reiterate the comments here as follow:

It can be mathematically proven that the steepness parameter *h* is confounded to the S_0 or R_0: The original Beverton-Holt stock-recruitment model is

1). $R_t = \dfrac{S_t}{\alpha + \beta S_t}$. Let this be rewritten as:

2). $R(S_t) = \dfrac{S_t}{\alpha + \beta S_t}$ since recruitment is a function of the stock size, where the α is the productivity parameter, representing the number of recruits per spawner at low numbers of spawners where the slope at the origin is $1/\alpha$. β controls the level of density dependence.

This definition was re-parameterized by a so-called steepness parameter, *h,* and reference-point type parameters at virgin population. The steepness parameter *h* is defined as the ratio of R_t to R_0 when $S_t = 0.2S_0$. The steepness parameter can also be interpreted as the fraction of the number of recruits in the virgin population (i.e. at time t_0) that is attained when its breeding biomass at time *t* is 20% of the virgin breeding biomass. Mathematically *h* is defined as:

3). $h = \dfrac{R(0.2S_0)}{R(S_0)} = \dfrac{R(0.2S_0)}{R_0}$.

It can be seen intuitively that with this re-parameterization, the parameters, *h*, S_0 and R_0 are highly confounded, which has made the parameter estimation impossible. Mathematically, this can be proven as follows. If *h* is independent of S_0 or R_0 from equation 3), then let us generate equation 3) more as:

4). $h(\lambda) = \dfrac{R(\lambda S_0)}{R(S_0)} = \dfrac{R(\lambda S_0)}{R_0}$ for any λ.

Then when λ=0.2, we get back to our traditional *h*, i.e. $h = h(0.2)$.

Therefore by re-arranging equation 4) as $R(\lambda S_0) = h(\lambda)R(S_0)$, for any λ and μ,

5). $R(\lambda\mu S_0) = R([\lambda\mu]S_0) = h(\lambda\mu)R(S_0)$

and

6). $R(\lambda\mu S_0) = R(\lambda[\mu S_0]) = h(\lambda)R(\mu S_0) = h(\lambda)h(\mu)R(S_0)$

Since the left side of equation 5) equals to the left side of equation 6), both the right sides should be equal, which leads to:

7). $h(\lambda\mu)R(S_0) = h(\lambda)h(\mu)R(S_0)$

This implies $h(\lambda\mu) = h(\lambda)h(\mu)$. Mathematically, $h(\lambda)$ should have the form of $h(\lambda) = \lambda^{\kappa}$ where κ is a constant.

Then from equation 3), $R(0.2S_0) = 0.2^{\kappa} R(S_0) = R_0$, i.e. the recruits of $R(0.2S_0)$ is proportional to, R_0 and we know from BH model this is not true. Therefore the definition of steepness parameter in equation 3) is not independent of R_0 and is instead a function of R_0, which cannot then be estimated independently.

Another way to eliminate this confounding effect is to re-parameterize the SR using the management parameters as illustrated in Schnute and Kronlund (1996, 2002) or Schnute and Richards (1998).

References

Chen, D.G. (2001). Detecting Environmental Regimes in Fish Stock-Recruitment Relationships by Fuzzy Logic. Canadian Journal of Fishery and Aquatic Sciences, 58(11):2139-2148.

Methot, Richard. 2006. User Manual for the Assessment Program Stock Synthesis 2 (SS2). Model Version 1.21. January 20, 2006. .NOAA Fisheries Seattle, WA.

Millar, R. B. (2004). Sensitivity of Bayes estimators to hyper-parameters with an application to maximum yield from fisheries. Biometrics. 60(2):536-542.

Schnute, J. T. and Kronlund, A. R. (1996). A management oriented approach to stock recruitment analysis. *Can. J. Fish. Aquat. Sci.*, 53:1281-1293.

Schnute, J. T. and Kronlund, A. R. (2002). Estimating salmon stock-recruitment relationships from catch and escapement data. *Can. J. Fish. Aquat. Sci.*, 59:433-449.

Schnute, J. T. and Richards, L. J. (1998). Analytical models for fishery reference points. *Can. J. Fish. Aquat. Sci.*, 55:515-528.

Appendix 1: List of Participants

Anne Hollowed

Delsa Anderl

Janet Duffy-Anderson

Craig Castelle

Peter Munro

Jennifer Ferdinand

Paul Spencer

William Stockhausen

Jack Turnock

Jim Ianelli

Tom Wilderbuer

Dan Nichol

Kerim Aydin

Appendix 2: Agenda of Flatfish review

CIE review, June 11-15th, 2007

Agenda (pdf version)

Monday June 11th			
9:00	Welcome and Introductions, adopt agenda,	**CIE Statement of work**	
9:15	Overview (species, surveys, fishery, catch levels, ABCs, TACs, bycatch)	Overview presentation	**Tom**
10:00	Biology (growth, natural mortality, diets, spawning & nursery areas, maturity)		
		Stockhausen presentation	**Buck**
		Duffy-Anderson presentation	**Janet**
		Nichol presentation	**Dan**
11:00	Age Determination of flatfish	Age and growth	**Delsa**
		Age validation	**Craig**
11:30	Trawl experiments on herding, escapement, bottom contact issues	Flatfish survey efficiency presentation	**Peter Munro**
12:00	***Lunch***		
13:00	Observer Program	N. Pac obs sampling	**Jennifer Ferdinand**
14:00	Summary of ongoing research	Duffy-Anderson presentation 2	**Janet**
15:00	Harvest control rules and projection model	(CIE Presentation)	**Jim**
Tuesday June 12th			
9:00	Gulf of Alaska and Bering Sea stock assessments		
	Overview of assessments	(CIE Presentation)	**Jim**
	BSAI yellowfin sole/northern rock sole	(CIE presentation)	**Tom**
	GOA arrowtooth flounder	(CIE presentation)	**Jack**
	BSAI Greenland turbot	(CIE presentation)	**Jim**
	GOA rex sole	(CIE presentation)	**Buck**
	GOA shallow-water flatfish	(CIE presentation)	**Buck**
12:00	**Lunch**		
13:00	**Incorporating ecosystem indicators in stock assessments**		
	Modeling catchability with bottom temperature	(CIE presentation)	**Tom, Paul**
	Catchability changes with temperature	(CIE presentation)	**Paul**
	Distribution changes with temperature	(CIE presentation)	**Paul**
Wednesday June 13th			
9:00	OSCURS model to define productivity regimes	(CIE Presentation)	**Tom**
	IBM 3-D drift model	(CIE presentation)	**Buck**
	Harvest control rules using F_{msy} for yellowfin sole and northern rock sole compared to proxy values	(CIE presentation)	**Tom**
12:00	**Lunch**		
13:00	Ecosystem aspects of flatfish	(CIE presentation)	**Kerim**
Thursday June 14th			
	Reviewer discussions with assessment authors		

Other background documents

List of Participants

SSC December 2006 minutes

Plan Team minutes and summary (2004 and 2005)

SAFE report Chapters and Intro sections

SAFE report checklist

SS2 Presentation (Methot)

Link to AFSC Survey data page

NMFS closed areas

Multispecies flatfish yield (Spencer et al. 2002)

Miller et al 2007 Catch estimation methods

Link to Estimates from Assessments

CIE report website

Appendix 3: Detailed Stock Assessment Review Activity

Monday, June 11, 2007

The meeting convened at 9:00am on June 11. Patricia Livingston, the Director of Resource Ecology and Fisheries Management Division, welcomed the group. Tom Wilderbuer, the review meeting coordinator, then opened the meeting with an overview of the flatfish fisheries and management which included the species with the management categories, the fishing areas, catches, relative stock sizes and the council/assessment process with the management tier system.

Buck Stockhausen, Janet Duffy-Anderson and Dan Nichol then presented the flatfish biology which included the growth, natural mortality, diets, spawning, nursery areas and maturity. Specifically:

- Buck Stockhausen presented an overview of flatfish biology and ecology in the Alaska region. Specifically, he talked about the BSAI/GOA geography, ecosystem, species diversity, climate variability and flatfish recruitment, and species information.
- Janet Duffy-Anderson talked about the dispersal of North Pacific flatfish larvae with outlines of recruitment dependent on larval transport to favorable nursery habitat, pelagic larvae dispersed to nursery habitats where juveniles assume a demersal existence, coupling between dispersal and development, sensitive to oceanographic changes and Sentinels.
- Dan Nichol presented the maturity/growth, spawning information for the flatfish species and migration behavior of Northern Rock sole.

The meeting proceeded to age determination. Delsa Anderl presented the Flatfish Age program at AFSC. A general ageing protocol on age determination and validation, and corresponding specifics for flatfish were nicely presented. After Delsa's general presentation on age determination, Craig Kastelle talked about age validation of Dover sole and yellowfin sole using bomb-produced radiocarbon (△C-14) to validate break-and-burn ages of individual Gulf of Alaska Dover sole and Bering Sea yellowfin sole. The conclusion is that the ages are accurate.

The morning meeting ended after the presentation from Peter Munro on flatfish trawl catchability to efficiency. After the lunch, the afternoon meeting re-convened and Jennifer Ferdinand gave a presentation on the North Pacific Groundfish Observer Program for its sampling strategies and methods. The Observer program had coverage from 200% to 0%. There were 200% coverage for vessels 125 ft and greater fishing pollock with two observers on board and all fishing events sampled. There were 100% coverage for vessels 125 ft and greater with one observer on board and random selection of fishing events sampled. There were 30% coverage (for 20% vessels) for vessels between 60 and 125 ft with 30% of fishing days sampled by random selection of fishing events. There was 0% coverage (for 72% vessels) for vessels under 60ft. The priorities for the observer program are to record fishing effort information for every haul (set) retrieved, sample for official total catch estimates and species composition, and send data to the observer program in Seattle, etc.

After this presentation, Janet Duffy-Anderson gave a summary of on-going research and the Monday meeting ended with the presentation by Jim Ianelli on the projection models for North Pacific Groundfish stocks.

Tuesday, June 12, 2007

The Tuesday meeting was mainly on flatfish stock assessment methodologies and to incorporate ecosystem indicators into assessment.

- **Jim Ianelli's model overview**

The morning meeting started with Jim's presentation on general overview on stock assessment methods. The general models are based on the age-structured model with the standard catch equation as the operational population dynamics model coded into ADMB to optimize the likelihood components. Stock-recruitment models are based on BH, Ricker and historical averages, and Bayesian priors were used for some parameters for posterior distribution.

- **Tom Wilderbuer presented yellowfin sole/rock sole**

Data Sources for the yellowfin sole are: 1) fishery catch 1954 – 2006; 2) fishery age compositions from 1964 to 2005, 3) survey age compositions in 1975 and from 1979 to 2005, 4) survey biomass from 1982 to 2006, 5) maturity and 6) weight at age from selected surveys, observer sampling. The age-structured model resulted in 131 parameters to be estimated from ADMB with Fishing mortality (52 parameters), Selectivity (4 parms), Survey catchability (2 parms), Year class strength (71 parms) and Spawner-recruit (2 parms). Model fits were satisfied and assessment results were appropriate.

For Northern rock sole, the data components included fishery and trawl survey age compositions, survey biomass and standard error, sexes are combined, selectivity is fixed asymptotic for older fish, Ricker form of the stock-recruitment curve is fit inside the model, and catchability (q) is constrained to a value near the estimate of q from a trawl herding experiment using the shelf survey trawl and natural mortality is estimated as a free parameter. Several sensitivity runs were conducted for sensitivity analysis for the natural mortality parameter M. In general, the model fits were appropriate even though the model does not fit the highest survey point estimates but the trend is fit.

- **Jack Turnock presented the assessment for GOA Arrowtooth flounder.**

For this stock, the data components included catch from 1960 to 2005, IPHC trawl survey biomass and SE from 1961 to 1962, NMFS exploratory research trawl survey biomass and SE from 1973 to 1976, NMFS triennial trawl survey biomass and SE from 1984 to 2005, fishery size components from 1977 to 1981, 1984 to 1993, 1995 to 2005, NMFS survey size compositions in 1975 and 2005, NMFS triennial trawl survey age composition from 1984 to 2003.

The stock assessment model had a split sexes, survey selectivity is estimated as asymptotic, female natural mortality is fixed at 0.2, males are estimated at 0.35 from profiling over a range of values and evaluation of model results, length data fit using a transition matrix of age to length, 50% Maturity at 47 cm, age 7.5.

The model fits are generally appropriate.

- **Jim Ianelli on Greenland turbot**

Assessment data used for this stock are survey size at age data in 1994 and 1998, shelf survey size composition and biomass estimates from 1979 to 2006, slope survey size composition and biomass estimates in 1979, 81, 82, 85, 88, 91, 2002, 2004, longline survey size composition and abundance index from1996 to 2006, total fishery catch data from 1960 to 2006, trawl fishery size composition from 1977 to 87, 1989 to 91, 1993 to 2006, longline fishery size composition in 1977, from 1979 to 85, 1992 to 2006.

The assessment was the SS2 which was, for the first time, used for this stock. The CIE panel recommended that the author investigate the appropriateness of the re-parameterization of the SR model to the steepness formulation as the application to this stock.

The author presented the same results from the conventional tagging studies for this stock and it did show a migration pattern between Russia stocks to the US side. The CIE panel suggested that a co-op assessment might help.

The CIE panel also noticed and argued against the appropriateness of the ABC-catch recommendation based on the sharp decline of the estimated abundance and biomass.

- **Jack Turnock on Rex Sole**

Similar data as arrowtooth stock were used in the assessment of this stock, which included fishery catch data from 1982 to 2005, fishery length comps from 1982 to 84 and from 1990 to 2005, survey biomass estimates triennially in 1984-1999, 2001, 2003, 2005, survey length comps in 1999, 2001, 2003,2005, survey age comps in 1984, 1987, 1990, 1993, 1996.

The assessment model is an age-structured model with 74 estimated parameters including selectivity curve parameters (fishery & survey, male & female), mean log recruitment and recruitment deviances, average fishing mortality and fishing mortality deviances, ages 3 – 20+, lengths 9 – 65+ cm (2-cm bins), catchability (q) fixed at 1.0, M estimated at 0.17 for both females and males from max age of 27.

The model fits were generally appropriate except some lack of fit for survey age-comp data. This is a tier 5 projection from tier 3 assessment. The CIE panel recommended that a Tier 3 assessment for this stock is appropriate.

- **Jack Turnock on GOA shallow-water flatfish (Tuesday afternoon)**

Several data sources including fisheries and survey data were presented and investigated. Fisheries data include regional total catches and observer data. Survey data were from the GOA groundfish trawl survey and was deemed reliable. Therefore, this stock was assessed as tier 4 or 5.

During the presentation, it was concluded that the other data were available, but not processed for the age-structured model. For future research and as a recommendation, the CIE panel recommended developing an age-structure model for this stock in the future.

- **Tom Wilderbuer Modeling catchability with bottom temperature**

Tom presented his discoveries on modeling catchability with bottom temperature with a log-linear model as $q = e^{\alpha+\beta T}$. The survey catchability (q) is estimated for each year in the model by considering the relationship to annual bottom water temperature as $B_{survey} = qS_a N_{a,y} w_a$, which improved the fit to the survey biomass.

- **Paul Spencer on catchability changes and distribution changes with temperature**

Following the same topic to model the catchability with temperature, Paul Spencer presented "Estimation of temperature-dependant catchability for eastern Bering Sea flatfish".

Paul Spencer then presented the distribution changes with temperature as "Geographic distributions of eastern Bering Sea flatfish: Effects of environmental variability and population abundance". The BS is a highly productive system with primary productivity ranges from 200 to 800 g C m^{-2} and sea ice plays a major role in primary and secondary productivity, particularly in the formation of the "cold pool". Strong statistical relationship between fish distribution and cold pool distribution has been found for several flatfish species.

Wed, June 13, 2007

- **Tom Wilderbuer on OSCURS model to define productivity regimes**

Tom talked about the use of the OSCURS model to define productivity regimes. OSCURS (Ocean Surface Current Simulation Model) model converts the selected daily sea level pressure grid to east-west wind velocity components and then current grids using empirical functions. The pressure grid includes most of the North Pacific Ocean from Baja California to China for all years since 1901. Details can be found at http://www.afsc.noaa.gov/REFM/docs/oscurs/get_to_know.htm. Model trajectories were compared with trajectories measured with satellite tracked drifters to calibrate the model. The model has been validated from predictions of the landfall (time and area) of spilled cargo from container ships. A 1988 regime was identified for rock sole and arrowtooth from their stock-recruitment data. The CIE panel recommended the assessment author verify whether this is a real regime shift since it has been debated in the literature.

- **Buck Stockhausen on IBM 3-D drift model**

Buck presented "Early Life History Stage Models for Dispersal and Recruitment of Flatfish" for combining early life history information and models of physical transport.

- **Tom Wilderbuer on "Harvest control rules using Fmsy for yellowfin sole and northern rock sole compared to proxy values".**

Generally in order to determine the preferred harvest strategy, it is important to know whether or not the spawner-recruit curves estimated for yellowfin sole and rock sole are reliable (are they a true representation of the productivity of the stock). A series of simulations was done on whether to include the pre-1977 SR data. It was shown that a Ricker model to the data from time frame (1978-2000) appears very reasonable and therefore FMSY is well estimated for yellowfin sole. Similarly, if using data from 1978 to 2001, Fmsy and Bmsy reasonably well estimated for Rock sole. The CIE panel recommended investigating density dependence and climatic changes in SR.

- **Kerim Aydin on Ecosystem aspects of flatfish**

In the afternoon, Kerim Aydin presented "Multispecies/ecosystem modeling with reference to flatfish" in the development of a conceptual model of the food web. This MS model would have advantages to improve system for non-target species management, to incorporate ecosystem considerations into individual stock assessments, to define ecosystem level reference points, etc. The CIE commented on this work and suggested for the future modeling to incorporate habitat, climatic data, competition, disease, etc., and ecosystem interaction/food interaction.

Thursday, June 14, 2007

The CIE panel spent the morning exchanging ideas and discussing with assessment authors. The meeting adjourned after lunch.

Appendix 4: Consulting Agreement between Dr. Din Chen and NTVI

Alaska Flatfish Review Statement of Work

General

The Alaska Fisheries Science Center (AFSC) requests a peer review of the Gulf of Alaska and Bering Sea/Aleutian Islands flatfish stock assessments by the Center for Independent Experts (CIE). Flatfish fisheries in Alaska are valuable and are likely to receive increased interest from commercial fishers in the next decade. Most flatfish populations are currently at stable or high levels of abundance. The North Pacific Fisheries Management Council (Council) is likely to pass amendments to rationalize (establish quota shares) groundfish fisheries in the near future. These amendments will allow for more flexibility in the time and areas fished which may result in the development of renewed interest in flatfish fisheries. Thus the AFSC desires an independent peer review of these stocks to assess the quality of the assessments and to determine whether the Council is being provided with the best available information and analysis.

There are currently 21 flatfish species managed under the Bering Sea and Aleutian Islands Fishery Management Plan, of which six species have age-structured stock assessments. In the Gulf of Alaska there are ten managed species, of which four have age-structured assessments. These assessments have never undergone a CIE review. While the current Council review process is in place, there remains a compelling need for an independent peer review of the Alaska flatfish assessments.

The CIE assessment review requires a total of four reviewers who are thoroughly familiar with various subject areas involved in stock assessment, including population dynamics, separable age-structured models, harvest strategies, survey methodology, and the AD Model Builder programming language. They should also have experience conducting stock assessments for fisheries management. The reviewers will travel to Seattle, Washington, to discuss the stock assessments with the flatfish assessment authors and other scientists at the AFSC involved in flatfish stock assessment. Overview presentations will be made on the topics listed below and assessment authors will be available for questions from reviewers. Specific details on a few (3-4) selected stock assessments will be presented to assist in the review.

Three of the reviewers shall generate individual reports. The fourth reviewer shall generate a Summary Report that compiles the points made by the three individual reviewers into one succinct document. The individual reports shall be appended to the Summary Report, thereby providing the complete detailed information from the individual reviewers. All reports shall address the strengths and weaknesses of the following points.

- Modeling efforts for Bering Sea/Aleutian Islands and Gulf of Alaska flatfish assessments and harvest recommendations. Specifically, the review shall evaluate:

  - The analysts' use of fishery dependent and fishery independent data sources in the assessments;
  - Gaps or inconsistencies in the population dynamics modeling methodology or logic;
  - If uncertainties in assessment model results are appropriately applied to management advice; and
  - Whether the assessments provide the best available science.

- The effort to incorporate ecosystem indicators and shifts in states of nature in the assessments. These include modeling survey catchability with annual bottom water temperature and using the Ocean Surface Current Simulation Models (OSCURS) to define putative oceanic productivity regimes.

- The harvest control rules adopted for Bering Sea yellowfin sole and northern rock sole (where a stock-recruitment model and F_{MSY} quantities are estimated) compared to other flatfish stocks where proxy values are used. Specifically, comments on the trade-offs between the different approaches are required.

The AFSC will provide copies of stock assessment documents, survey reports, and other pertinent literature on a web site.

Specific

The CIE shall provide four reviewers, for a maximum total of 52 work days. The three individual reviewers are approved for a maximum of 15 work days each, and the summarizer is approved for a maximum of seven work days. The three individual reviewers shall be approved for travel to the meetings at the AFSC. The summarizer shall work from his or her home office, so no travel is required.

Specific requirements for the individual reviewers and the summarizer are listed below[2]. The list is followed by a chronological table, which includes the required steps for the CIE. If any intermediate step in the review process is delayed, all subsequent steps that depend on it will be delayed by an equivalent period.

<u>Individual Reviewers</u>

1. Read and become familiar with the relevant documents provided to the reviewers. The ten age-structured assessments that encompass the focus of the review are presented in 512 pages (of which approximately 1/3 is text).
2. Discuss the stock assessment with the lead assessment scientist and survey scientists at the AFSC, in Seattle, Washington, from June 11 to June 14, 2007 (see attached agenda).
3. No later than June 29, 2007, submit a written report of findings, analysis, and conclusions. More details on the report outline and organization are provided in

[2] All reports will undergo an internal CIE review before they are considered final.

Annex I. The report shall be sent via e-mail to Dr. David Die (ddie@rsmas.miami.edu), and to Mr. Manoj Shivlani (mshivlani@rsmas.miami.edu), and to the CIE summarizer (contact information to be provided by the CIE).

Summarizer

4. Become familiar with the documents provided to the individual reviewers.

5. Read the draft reports from the individual CIE reviewers, and draft the summary report following the standard outline. More details on the report outline and organization are provided in Annex I.

6. Complete summary report within one week of receiving final individual reviewer reports from the CIE. The report shall be sent via e-mail to Dr. David Die (ddie@rsmas.miami.edu), and to Mr. Manoj Shivlani (mshivlani@rsmas.miami.edu).

Activity	**Deadline**
Individual reviewers submit their reports to the CIE and the summarizer.	June 29, 2007
CIE reviews and approves the individual reviewer reports, and provides them to the NMFS COTR.	July 13, 2007
NMFS COTR approves individual reviewer reports.	July 17, 2007
CIE provides final individual reviewer reports to NMFS COTR and to the summarizer.	July 20, 2007
Summarizer provides summary report to CIE.	July 27, 2007

CIE reviews and approves the summary report, and provides it to the NMFS COTR.	August 10, 2007
NMFS COTR approves summary report.	August 14, 2007
CIE provides final summary report to NMFS COTR.	August 17, 2007

Submission and Acceptance of CIE Reports

The CIE shall provide the final reports for review for compliance with this Statement of Work and approval by NOAA Fisheries to the COTR, Dr. Stephen K. Brown (Stephen.K.Brown@noaa.gov), according to the above schedule. The COTR shall notify the CIE via e-mail regarding acceptance of each report. Following the COTR's approval, the CIE shall provide a pdf format version of each approved report to the COTR.

Annex 1: Contents of Reviewer Reports

The following requirements refer to all reports, both the individual reports and the Summary Report.

1. All reports shall be prefaced with an executive summary of findings and/or recommendations.

2. The main body of all reports shall consist of a background, description of review activities, summary of findings, conclusions/recommendations, and references.

3. The reports shall also include as separate appendices the bibliography of all materials provided and any additional papers cited, along with a copy of the statement of work.

Summary report only: The summary report shall include the three individual reviewer reports as appendices.

Review by Dr. Paul Medley

Independent System for Peer Reviews
Consultant Report on:

Alaska Flatfish Stock Assessments

June 11 - 14, 2007

Paul A. Medley
Sunny View
Jack Hole
Alne, YO61 1RT
UK
paul.medley@virgin.net

Executive Summary

The analysts' use of fishery dependent and fishery independent data sources in the assessments is appropriate and well-founded. There are some potential problems with data sources, but they are probably not significant and out of the control of the AFSC scientists. For example, the abundance surveys are not designed specifically for flatfish, but only take flatfish species as bycatch.

There are no gaps or inconsistencies in the population dynamics modelling methodology or logic, although some improvements have been suggested. It has been found that these suggested changes are unlikely to make any qualitative difference to the assessments. However, it is good practice to use research which is complete, subject to review and thought to best represent processes in the population dynamics. Where appropriate, it has been suggested these be incorporated into the assessment and projection models.

The uncertainties in the assessment model are represented in the management advice. There are improvements in representing structural uncertainties in the model which could be undertaken in future. In particular, further development of the management strategy evaluation (MSE) approach can be used to identify important uncertainties and the research required to resolve them.

The assessments apply the best available science. The stock assessment and associated research is of high quality. The AFSC scientists are applying best practice and there is clear evidence of on-going improvements in the assessments and understanding of the population dynamics.

The approach incorporating ecosystem indicators and environmental effects in the assessments is well done and should produce better assessments over time. However, it is likely that work on climate shifts and the ecosystem can only be fully included in management advice through using a MSE approach.

It is appropriate to apply the tier based harvest control rules as recommended by the SSC for yellowfin sole and northern rock sole and other species as appropriate. While the AFSC scientists should be and are concerned about model uncertainties, the on-going monitoring and the way the management controls are applied should make the system robust to such uncertainties. Management can be tested and improved through carrying out further work on management strategy evaluations.

Introduction

The Alaska Fisheries Science Center (AFSC) requested a peer review of the Gulf of Alaska and Bering Sea/Aleutian Islands flatfish stock assessments. There are currently 21 flatfish species managed under the Bering Sea and Aleutian Islands (BSAI) Fishery Management Plan, of which six species have age-structured stock assessments. In the Gulf of Alaska (GOA) there are ten managed species, of which four have age-structured assessments.

A stock assessment and fishery evaluation (SAFE) report is prepared and reviewed annually for each fishery management plan (FMP). The SAFE report summarizes the best available scientific information concerning the condition of the stocks, marine ecosystems, and fisheries, which the Councils use for determining annual harvest levels among other things.

The Stock Assessment sections of the SAFE reports for the groundfish fisheries are compiled by the BSAI and GOA Groundfish Plan Teams from chapters written by NMFS Alaska Fisheries Science Center (AFSC) scientists. These chapters include a recommendation for the overfishing level (OFL) and acceptable biological catch (ABC) for each stock and stock complex managed under the FMP. The ABC recommendations are reviewed by the Scientific and Statistical Committee (SSC), which may confirm the Plan Team recommendations or develop its own. The ABC recommendations, together with social and economic factors, are considered by the North Pacific Fishery Management Council (Council) in determining the total allowable catches (TACs) and other measures used to manage the fisheries.

The stock assessment section of the flatfish SAFE reports is the subject of this review. The reviewers (Din Chen, Paul Medley, Graham Pilling) attended a meeting in Seattle, Washington 11-14th June 2007, to discuss the stock assessments with the flatfish assessment authors and other scientists at the AFSC involved in flatfish stock assessment. Overview presentations were made on the assessments and associated topics. The assessment authors were very helpful answering questions and supporting the review.

Of the 21 species, the main stock assessments covered by this review were arrowtooth flounder (*Atheresthes stomias*) and rex sole (*Errex zachirus*) in the Gulf of Alaska, and Greenland turbot (*Reinhardtius hippoglossoides*), northern rock sole (*Lepidopsetta polyxystra*) and yellowfin sole (*Limanda aspera*) in the Bering Sea and Aleutian Islands. Comments on these stock assessments are made within the overall comments on the different aspects of the assessments. As there were many stock assessments to consider, the individual stock assessments were not reviewed in detail, but the general strategy and approach to flatfish stock assessment was covered with attention paid to those issues highlighted by the AFSC scientists, SSC and Plan Teams.

Summary of Findings

Fishery Independent Data: Abundance Indices

The assessments rely mainly on bottom trawl surveys, although sablefish longline surveys are also used in the Greenland turbot assessment. The bottom trawl survey is not specific to flatfish, raising concerns that the surveys may miss significant biomass, be noisy and subject to climatic and other effects. Nevertheless, the accurate ageing together with the survey data seems to generate indices of abundance accurate enough for stock assessment.

The surveys are used to estimate year class sizes within the catch-at-age models. Simple catch curves can be used to indicate the average levels of total mortality on each year class. Although there are a few outliers, the overall results suggest the yellowfin sole data are reliable, and mortality can be estimated with a reasonable degree of confidence (Fig. 1). Catch curves are much cruder than the fitted catch-at-age model as they take no account of changes in the level of exploitation, but use the same contrast pattern in the data to estimate mortality. The total mortality estimates suggest that Z has been about 0.17 year^{-1}, which after allowing for M suggests fishing mortality is low for yellowfin sole, agreeing with the assessment.

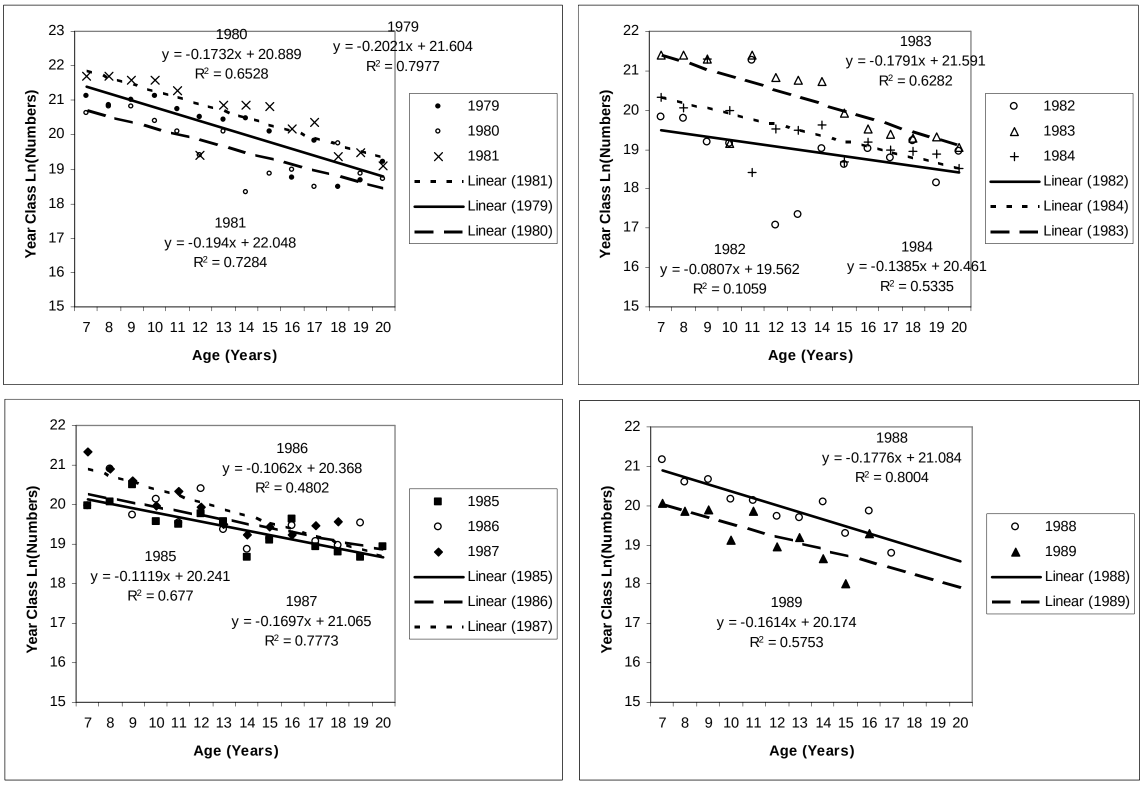

Figure 1 Catch curves for yellowfin sole for the year classes 1979-1989 based on survey data from the age of 7 when selectivity remained constant. A similar pattern can be observed for northern rock sole, although the average total mortality is about twice as high (0.3 year^{-1}). Although there are a few outliers (e.g. 1982 year class), the pattern is fairly consistent from year class to year class.

The research on trawl efficiency, the herding effect and standardisation using bottom temperature and types of escapement from the survey net should significantly improve the survey indices. The exponential model used to describe the relationship between temperature and catchability is a good choice if the underlying cause of the catchability change is due to physiology. The physiological response (e.g. swimming activity) is likely to increase in an exponentially with temperature until other limiting factors, such as the availability of oxygen, take effect. Bottom temperature standardisation should be carried out for all the flatfish species.

The log-normal is not always the most appropriate likelihood. For the trawl surveys the log-normal is probably a reasonable choice. Accumulated trawl effort will rarely generate 0 catch and due to the aggregated nature of populations the variance is likely to increase with abundance. However, as part of the documentation of model diagnostics, this should be checked, by, for example, plotting the residuals against the expected survey index values.

The log-normal is not necessarily a good choice for longline surveys, however. There are good theoretical reasons for choosing the binomial, Poisson negative binomial or beta-binomial as appropriate likelihoods for hook based gear. In terms of fitting, the practical result is to weight indices in favour of low values. The two recent years' longline survey indices for Greenland turbot show low values, suggesting that Greenland turbot abundance has decreased in recent years, in contrast to the trawl survey which suggests an increase in response to lower catches. The narrow confidence intervals on the recent years' indices could be an artefact of the choice of likelihood and underestimate the uncertainty. Therefore assessment scientists should consider an alternative likelihood for the longline survey.

Fishery Dependent Data: Catch Data

The catch data appears to be reliable and well estimated. With the observer coverage, discarding is probably well estimated. There is concern that there may be an "observer effect", so that discarding is higher when observers are not on board. It may be possible to look at vessel monitoring system information, as well as use port sampling to check whether there is a significant change in vessel behaviour at sea or change in what is landed when an observer is on board. This issue does not threaten the validity of the assessments.

The stock assessment most affected by possible errors in catch data is Greenland turbot. The Greenland turbot is probably the least reliable stock assessment, but most heavily exploited stock. There is a potential problem with the assessment if a significant amount of catch is taken outside USA jurisdiction in Russian waters. This would also affect the management of this species. It appears from trawl surveys that the amount of biomass of other flatfish species in Russian waters is small, so that, in contrast to Greenland turbot, these can be safely treated as a management unit. Greenland turbot also has less age data than other flatfish species, so there is greater reliance on size composition data from which age is inferred. This introduces greater uncertainty in the assessment.

Interpretation of the 1960-1980 period is important in determining the state of the Greenland turbot stock. The key period 1960-1980 is not covered by the abundance indices or other data apart from catches. It is not possible to directly estimate recruitment in these years and a stock recruitment relationship has to be assumed.

While the current SS2 model does explain fluctuations in abundance over 1980-2006 well, these are minor compared to the overall changes which have occurred since the fishery began. To illustrate the problem, a simple biomass dynamics model (Schaefer) was fitted to the catch and survey index data. The biomass dynamics model is crude, assumes a simple direct biomass-recruitment relationship and takes no account of the available age and size composition data. The estimate for the rate of increase in this model is probably too low (it should be around double the natural mortality, but is estimated as less) making the assessment pessimistic, and it does not explain the abundance indices as well as the age structure model can. However, the biomass dynamics model does produce approximately the same trajectory of the biomass in recent years as the catch-at-age model. This model gives an alternative account of the biomass in the early years using the available catches (Fig. 2). Many early biomass trajectories are plausible, and any information to help determine 1960-1980 abundance would be very valuable.

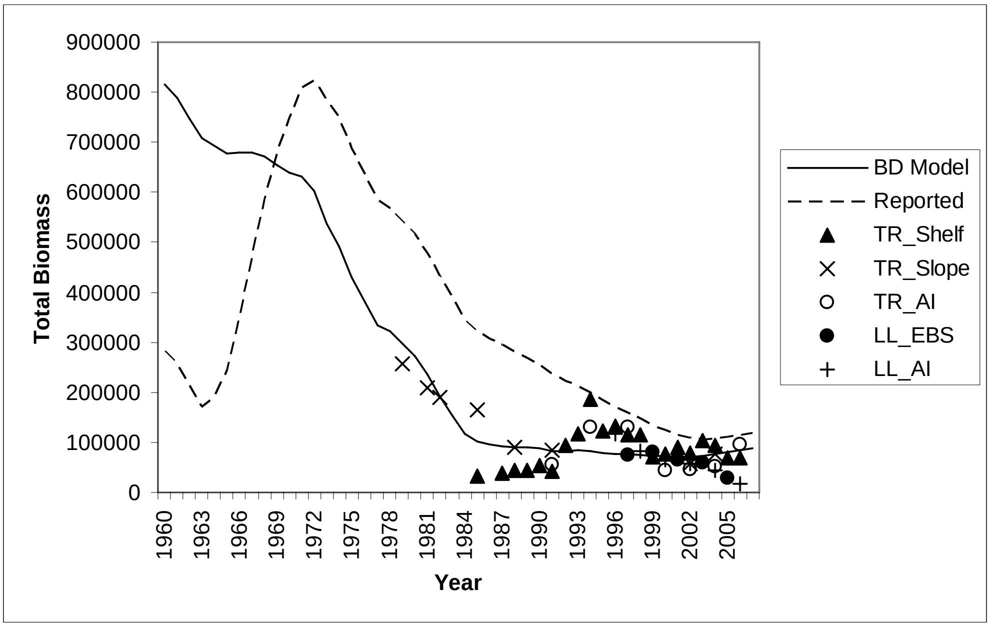

Figure 2 Estimated biomass and observed abundance indices for Greenland turbot. The "Reported" biomass is the biomass estimated for the stock (SAFE, 2006). The "BD Model" biomass was estimated using a simple biomass dynamics model fitted to the trawl (TR) and longline (LL) abundance indices. The early years are extrapolated using the models and catch data which is available back to 1960.

Ecosystem Modelling and Management Strategy Evaluation

Currently the ecological modelling has at least four aims:

- Improve estimates of natural mortality by accounting for predation. This has given some support for the choice of natural mortality in the assessments.
- Explain ecological patterns observed on the Alaska shelf ecosystem, such as increased species evenness and increase in arrowtooth flounder abundance.
- Improve the understanding of the ecosystem and its relationship to climate. Research currently being undertaken is appropriate to address this issue.
- Develop methods for estimating the optimal yield from fish guilds.

While the current Ecopath-with-Ecosim (EwE), modelling will be able to describe the ecosystem in some detail, it is likely that the models will eventually fall short of what is required, and it will be necessary to move to a more detailed process model (e.g. "Atlantis" and "In Vitro" modelling approaches used in Australia; McDonald *et al.* 2005). This type of model describes many processes explicitly and includes the representation of the impacts of a wider set of human sectors and each facet of the adaptive management cycle (biophysical system, human activity sectors, monitoring, assessment and management activities). Such models can include an individual-based models (such as the IBM dispersal model) and are a flexible tool to develop an operational model that best represents how it is believed that the ecosystem functions and interacts with human activities. While these models are much more complex to implement than EwE, they will make better use of current and future research. Much of the preparatory work on, for example, current and other oceanographic processes has already been undertaken.

An important feature of detailed process ecosystem models would be the ability to apply a management strategy evaluation (MSE; Sainsbury *et al.* 2000). In particular, the model should include: multiple representations of the main ecological groups and processes found in marine ecosystems; fleet-level to individual-level representation of the dynamics of human sectors (primarily fisheries); a sampling model that reports fisheries statistics and the results of fisheries independent data collection exercises, with error structures typical of that found in reality; assessment models (from simple estimates to the explicit assessment models used in reality; biomass dynamics models, catch-at age models etc); all major management controls (e.g. gear, effort, zoning, seasonal closures, quotas and the like) and management decision models, including lag in implementation of decisions; and simple to complex representations of the socio-economic drivers and considerations (particularly for compliance).

Even if AFSC plans to develop their own model, it may be useful to review “Atlantis” and other models to see whether they can be adapted for Alaska, thereby avoiding repeating the same work unnecessarily. For example, the computer code for Atlantis is freely available and could be used as a base to develop dedicated software rather than programming from scratch.

The development of the management strategy evaluation for testing methods and approaches is an important step forward. It will be worthwhile continuing to improve the MSE approach, ultimately allowing more ecological information to be included in management without explicitly modelling the ecology in the stock assessment. However, the MSE approach will require improvements in the use of performance indicators to assess different scenarios efficiently.

Population Dynamics Modelling

The work presented at the meeting on understanding larval life history and distribution, and the possible connections between climatic effects and recruitment strength, was clearly valuable. The availability of current transport model and other oceanographic data and models make research in this area valuable and cost effective. However, this is likely to prove a difficult area to build reliable quantitative relationships as there will probably be a number of mutually dependent non-linear factors such as ice shelf extent, distribution of habitat, primary production as well as bottom temperature determining dynamics and availability of fish to the fishery.

A recruitment index would be valuable to these fisheries. A recruitment index of abundance after any density dependent mortality can be used to improve projections and short term forecasts. Catch-at-age models often predict poorly because they are best at estimating the size of the cohort once most of it has been caught. Recruitment indices help reduce this problem giving an indication of the abundance of year classes before they enter the fishery. An index of juvenile abundance is more likely to work as a recruitment index as it is less likely to be affected by large random mortalities and density dependence.

Monitoring juveniles and adults before maturity may allow consideration of the effects of climate on growth. The excellent ageing data for many of the flat fish species makes the detailed study of growth at this level of detail a possibility. It may be possible to make better use of otolith data, if the otolith width can be linked to length for some species, thereby obtaining an estimate of the growth history. Growth rate may be linked to mortality for the Lorenzen relationship (Lorenzen 2005), which may become a significant factor in juveniles.

Growth rates might also be affected by temperature and/or productivity. Allowing growth rates to change would probably make little difference to the population dynamics of older animals. However, it could still prove important in stages prior to maturity and in particular juveniles, if these are modelled in the future.

It makes sense to split the population dynamics models for the separate sexes where they differ in size. This does require more parameters, but the changing sex ratio and different size between males and females for many species implies that the life histories will be different.

The observed sex ratio might be explained to some extent by the size difference. The proportion male will depend on the relative cumulative mortality between the two sexes.

$$P_{ma} = \frac{R_m e^{-\sum_{i=0}^{a}(M_{mi}+F_{mi})}}{R_m e^{-\sum_{i=0}^{a}(M_{mi}+F_{mi})} + R_f e^{-\sum_{i=0}^{a}(M_{fi}+F_{fi})}}$$

where M_{mi}, M_{fi}, F_{mi}, F_{fi} =natural and fishing mortality for males and females at age i, and R_m, R_f = recruitment for males and females. Where the sex difference between the fishing mortalities is negligible and $R_m = R_f$, this simplifies to a logistic type model:

$$P_{ma} = \frac{e^{-M_m a}}{e^{-M_m a} + e^{-M_f a}}$$

where P_{m0} = 0.5 (sex ratios are equal) at the start age (0). It is likely that selectivities will apply to different sexes, perhaps related to length, making this a poor description of a heavily exploited population as the fishing mortalities (F_{mi}, F_{fi}) would be significant. Differences in abundance by sex might be explained by changing natural mortality at length (Lorenzen 2005):

$$M\ a = -Ln\left(\frac{L_0}{L_0 + L_\infty\left(Exp(k\ a) - 1\right)}\right)\left(\frac{M_1}{L_\infty k}\right)$$

where L_0 = length at the start of the period, a = the period over which the mortality is applied, M_1 = the natural mortality rate at unit length, and L_∞ and k are the von Bertalanffy growth parameters. Using this model, the same natural mortality at unit length can produce different natural mortalities for different sexes if their growth varies, but it was found this could not explain observed sex ratio in arrowtooth flounder (Fig. 3).

Alternatively, different mortality rates could be applied to the two sexes as was done in the original assessments where appropriate. The arrowtooth flounder female natural mortality was fixed at approximately 0.2 year^{-1} (M_1 = 11.0 for the Lorenzen model) and the male natural mortality was fitted to the sex ratio using least squares (Fig. 3). The Lorenzen model with separate sex mortalities fitted the sex ratio only slightly better than the fixed mortality model used in the assessment. Fishing mortalities and selectivity may be different between the sexes, which can only be explored as part of the full stock assessment model.

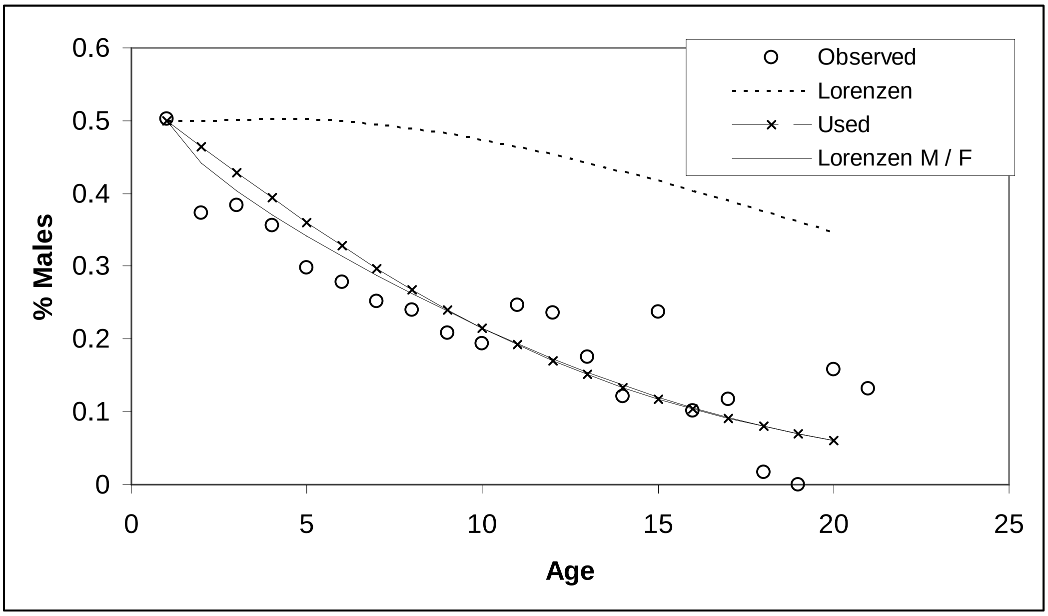

Figure 3 Sex ratio for arrowtooth flounder observed and predicted for the different model approaches. Lorenzen model with a single parameter for males and females, the fixed mortality model with separate male and female mortalities (used) and the Lorenzen model with separate male and female mortalities. To explain the observed sex ratios separate parameters are required for males and females.

The difference between males and female abundance suggests either a different mortality rate between sexes, probably related to different life histories, or different ratios of female and male eggs produced. The Lorenzen model with different natural mortality for each species may fit the observations a little better than the model used, but the difference is negligible.

Another observed pattern, at least in northern rock sole, is the negative correlation between growth rate and population size (Walters and Wilderbuer 2000). While past changes can continue to be described empirically using observed means, this makes the predictive power of the assessment and projections limited. If a process is thought to exist in the population dynamics, it probably should be included in the population models if possible. Based on experiences in aquaculture Lorenzen and Enberg (2002) suggest a linear relationship between biomass and maximum mean size:

$$L_{b\infty} = L_{\infty} - gB_t$$

While this model seems to represent a good description of the relationship over a reasonable range, it could present problems when fitting as negative values are possible. Non-linear functions can appear linear over a range of values, but should behave better at the extremes and provide more stability when fitting to data. One option might be:

$$L_{b\infty} = L_{\infty} e^{-g\,B_t/B_{\infty}}$$

Where B_{∞} = the unexploited biomass so that B_t/B_{∞} represents the state of the stock relative to the unexploited and therefore e^{-g} = proportional decrease of the maximum length possible due to density dependence. Incorporating this or similar model into relevant assessments and projections should improve their accuracy.

In the yellowfin sole stock assessment, SSB is calculated as equal to yellowfin sole weight multiplied by the proportion mature, which assumes fecundity is proportional to weight. The exponent for the weight length relationship is 3.06, whereas the exponent relating length to fecundity estimated by Nichol and Acuna (2001) is 3.63, indicating that fecundity increases at a faster rate with length than the weight, making older, larger females relatively more important. It was demonstrated that the estimate agrees with other fecundity studies for this species. However, this alternative model for calculating SSB will probably make little difference to stock assessment results (Fig. 4). Nevertheless, as the fecundity work has been completed, and appears reliable, there is little reason to rely on a fecundity proxy (maturity * weight) when an estimate is available.

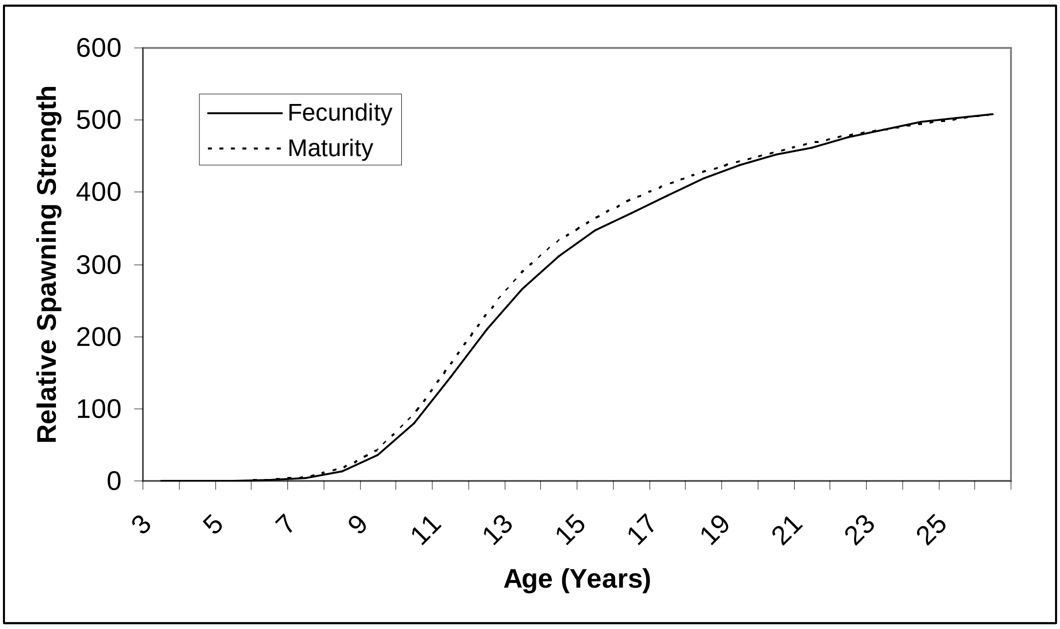

Figure 4 Relative spawning strength (scaled survival multiplied by maturity/fecundity ogive) for the maturity and fecundity models for yellowfin sole. The fecundity model shows a small increase in the relative importance of larger older females.

Uncertainties Applied to Management Advice

The uncertainties in the assessment model are represented in the management advice. Using the harvest rule tier system to represent risk should be avoided, but rather decision tables could be developed as a way to represent risk to managers. There are improvements in representing structural uncertainties in the model which could be undertaken in future. Further development of the management strategy evaluation (MSE) approach, described in the previous section, should be used to identify important uncertainties and the research required to resolve them.

Harvest Rules

It is beyond the scope of this review to comment on the tier system used for Alaska fisheries management. There was inadequate documentation available to give the full background and explanation behind the rules determining the tiers under which stocks are assessed. In general, the tiers are designed to allow greater levels of exploitation as information on the stocks is improved and the tier system appears to achieve this. However, there also appears to be a tendency to use the tier system to apply precaution in dealing with structural uncertainty, such as on the reliability of the stock recruitment relationship. If the tier system comes under detailed review, it might best be checked through a MSE approach.

The review was asked to consider specific harvest rules as they apply to the yellowfin sole, northern rock sole and rex sole stock assessments. The sensitivity analysis considered a number of sources of uncertainty for yellowfin sole, in particular comparing a single climatic period, so that the stock-recruitment relationship remains valid before and after 1978, and assuming two different climatic conditions before and after 1978, which is an appropriate comparison[3].

In the latter case, separate stock-recruitment relationships can be applied. However, on the basis that there may be two climate regimes, it was suggested by AFSC scientists to keep the species on a lower more precautionary tier. On this basis, however, it seems unlikely a species could ever be placed on tier 1. Given the reasonable fit of the stock-recruitment relationship, species should be promoted to management tier 1 harvest rules. Model uncertainty should be dealt with in other ways, by improving the method by which the precautionary and limit fishing mortality reference points and the OFL and ABC are separated, and by using other management controls. Appropriate management actions to reduce risk can be explored using management strategy evaluations.

Implementation of the flatfish harvest rules have not been tested in practice, because the halibut by-catch rules limit flatfish fishing mortality to a low level. It appears this will continue to be the case unless more selective gears are used.

---

[3] The SSC suggested that a more appropriate contrast between productivity regimes would be between the pre- and post-1978 datasets rather than between the full dataset and the post 1978 dataset. However, I believe that the AFSC scientists were comparing two hypotheses and their comparison was appropriate.

It is not clear the extent to which the method separating the OFL and ABC is precautionary. It may be that the current difference is based upon a general quadratic utility function. This type of function is used more for mathematical convenience than because it is a good description of utility functions. It is usual to spend a lot effort on improving knowledge through stock assessment, much less on assessing the utility of outcomes, through for example, socio-economic studies. Considering the cost and benefits of strategies and outcomes needs either detailed work from key decision-makers (e.g. Keeney and Raiffa, 1993) or more pragmatic approaches often used by businesses (e.g. Hwang and Yoon, 1981). Determining appropriate performance indicators in a management strategy evaluation may help to improve the method to recommend the OFL and ABC.

The fishing mortality reference points assume selectivity will not change. Concern was expressed that with a high estimated target fishing mortality, selectivity might then be altered invalidating the reference point and allowing overfishing. Assuming constant recruitment (more precautionary than assuming increased recruitment as with the Ricker stock-recruitment relationship), a per-recruit approach can be used to consider the effect of selectivity changes.

Biomass-per-recruit can be calculated for each species as the weight-at-age multiplied by the survival. While market demand will also control the optimum size of individual fish based on processing and use, the age at which the biomass-per-recruit is maximised indicates the approximate optimum age to harvest. Selectivity, to maximise yield, should cover the maximum biomass. The logistic selectivity should increase on or just before the maximum biomass-per-recruit point.

The yield-per-recruit iso-yield line indicates the fishing mortality which will provide the same yield-per-recruit as selectivity changes. Where the species is not the target, selectivity may not be optimised. For yellowfin sole and northern rock sole, the selectivity is already close to the optimum and there will be little incentive to change even if these species are targeted (Fig. 5). Rex sole shows a different pattern, with selectivity focused on much older fish (Fig. 6).

The F reference points are set assuming constant selectivity. If selectivity is not subject to management control and can be altered by fishermen, then the reference points can be invalidated. However, the primary control is not fishing effort; F is not controlled directly, but by setting a TAC. Hence the fishing mortality averaged over age achieved by a TAC will not be the same if selectivity changes. In this case, if rex sole was the target species fishermen would most likely get the TAC by reducing the age at 50% selectivity rather than increasing F. These results suggest the TAC control will be robust to selectivity changes.

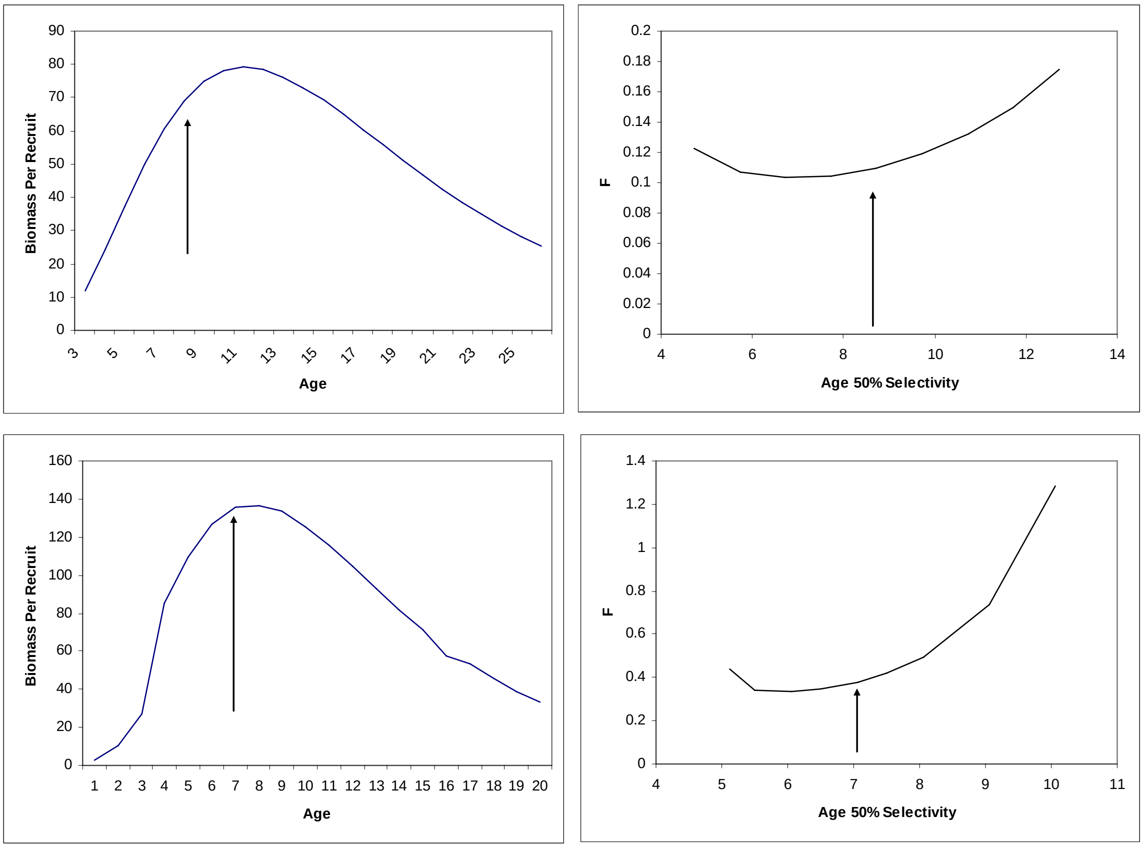

Figure 5 Yellowfin sole (top) and northern rocksole (bottom) biomass-per-recruit (left) and iso-yield line (right) with the current age at 50% selectivity (arrow). Significant changes in selectivity are unlikely as current selectivity is already close to the optimum.

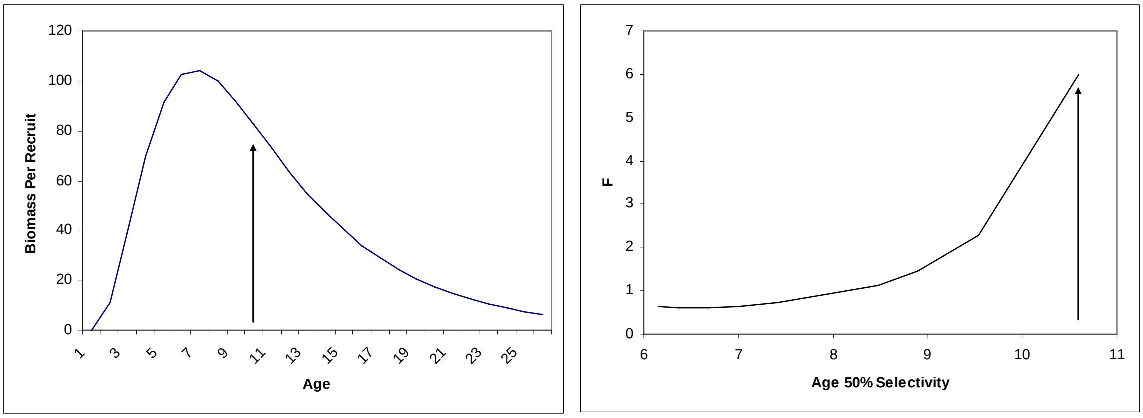

Figure 6 The current age at 50% selectivity (arrow) for rex sole is well to the right of the age when the cohort biomass would peak, suggesting that there would be a strong incentive for selectivity to shift should rex sole become a target species. Should selectivity change (and recruitment remain constant), the iso-yield line suggests that the TAC would be reached at a much lower F than the reference point.

Recommendations

The AFSC scientists should develop an abundance index based on the observer programme data. It should be possible to generate observer-based abundance indices for the flatfish to run alongside the current survey index. An index based on commercial data will require standardisation, using covariates such as trawl net type, depth, temperature, tow length, vessel speed and so on. Fishermen are likely to know what the important covariates affecting catchability are. The “observer effect”, where fishermen behaviour may be affected by the presence of an observer, will not matter as long as it has been consistent. It should be noted that a fishery dependent index is susceptible to management control changes.

The studies on survey catchability have been successful and should be continued. As well as improving the survey indices, this research will prove invaluable as the basis for standardising the observer data.

AFSC scientists should improve their communications and co-operation with Russian scientists, if possible, sharing survey and catch information. Combining catches may be particularly important in determining the state of the Greenland turbot stock. Unlike the other flatfish species, it is possible significant biomass is spread over a range outside Alaskan waters. Any historical data which could be used to develop an abundance index for Greenland turbot 1960-1980 would prove invaluable in the assessment of this species.

A process based ecosystem model should be developed with the intention of conducting more extensive management strategy evaluations. The MSE will need to consider how to measure performance of the management strategies being tested, through considering utility and regret functions, ecological and multispecies indicators as well as more traditional fisheries indicators and reference points. MSE are by their nature very complex, and will need to report simple results to allow effective evaluation. Indicators will need to be considered as much by managers and decision makers as by the scientists.

All models should, if possible, use retrospective analysis to test their accuracy. Retrospective analyses should be included in the stock assessment reports. An important test of stock assessments is to compare model predictions with outcomes, and models need to be regularly evaluated in this regard.

Where males and females show a marked difference in size, the sexes should be modelled separately. Although this requires more parameters, the changing sex ratio and different size between males and females implies that the life history strategies are different.

Significant population processes, such as density dependent growth if it is present, and more accurate estimates of population parameters such as fecundity, should be included in the assessments. Properly reviewed research should be used wherever possible, even if it makes little apparent difference. Updating assessments in this way should encourage further research and ensure the best science is included in the assessment.

References

Keeney, R.L. Raiffa, H. 1993. Decisions with Multiple Objectives. Preferences and Value Tradeoffs. Cambridge University Press, Cambridge, UK.

Lorenzen, K. & Enberg, K. 2002 Density-dependent growth as a key mechanism in the regulation of fish populations: evidence from among-population comparisons. *Proc. R. Soc. Lond. B* **269**, 49-54.

Lorenzen, K. (2005) Population dynamics and potential of fisheries stock enhancement: practical theory for assessment and policy analysis. Philosophical Transactions of the Royal Society of London. Fisheries Theme Issue 2004

McDonald, A.D, L.R. Little, R. Gray, E. Fulton, K.J. Sainsbury, and V.D. Lyne (2005) Multiple-use Management Strategy Evaluation for Coastal Marine Ecosystems Using InVitro. In Zerger, A. and Argent, R.M. (eds) MODSIM 2005 International Congress on Modelling and Simulation. Modelling and Simulation Society of Australia and New Zealand, December 2005, pp. 2588-2594. http://mssanz.org.au/modsim05/papers/mcdonald.pdf

Nichol, D.G. and Acuna E.I. 2001. Annual and batch fecundity of yellowfin sole, *Limanda aspera* in the eastern Bering Sea. Fish. Bull. 99:108–122 (2001).

Hwang, C.-L. and Yoon, K., 1981. Multiple Attribute Decision Making, Methods and Applications. A State of the Art Survey. Springer-Verlag, Berlin.

Sainsbury, K.J., Punt, A.E. and Smith, A.D.M. (2000) Design of operational management strategies for achieving fishery ecosystem objectives. *ICES Journal of Marine Science*, 57: 731–741

Walters, G.E. and Wilderbuer, T.K. 2000. Decreasing length at age in a rapidly expanding population of northern rock sole in the eastern Berings Sea and its effect on management advice. Journal of Sea Research 44: 17-26.

Appendix I: Document List

The following stock assessment chapters were made available through the AFSC web site:

Thomas K. Wilderbuer and Daniel G. Nichol. BSAI Northern Rock Sole

Thomas K. Wilderbuer and Daniel G. Nichol. BSAI Yellowfin Sole

James N. Ianelli, Thomas K. Wilderbuer, and Dan Nichol Assessment of Greenland Turbot in the Eastern Bering Sea and Aleutian Islands

Thomas K. Wilderbuer, Daniel G. Nichol and Paul D. Spencer Other Flatfish

Benjamin J. Turnock, Thomas K. Wilderbuer and Eric S. Brown Gulf of Alaska Arrowtooth Flounder Stock Assessment

Benjamin J. Turnock, Thomas K. Wilderbuer and Eric S. Brown Gulf of Alaska Flatfish

Benjamin J. Turnock and Z. Teresa A'mar NMFS Alaska Fisheries Science Center Gulf of Alaska Rex Sole Stock Assessment

Various other documents and data were made available through the AFSC website (http://www.afsc.noaa.gov).

Appendix II: Statement of Work

Peer Review of Alaska Flatfish Stock Assessments
June 11-14, 2007

General

The Alaska Fisheries Science Center (AFSC) requests a Center of Independent Experts (CIE) review of the Gulf of Alaska and Bering Sea/Aleutian Islands flatfish stock assessments. Flatfish fisheries in Alaska are valuable and are likely to receive increased interest from commercial fishers in the next decade. Most flatfish populations are currently at stable or high levels of abundance. The North Pacific Fisheries Management Council is likely to pass amendments to rationalize (establish quota shares) groundfish fisheries in the near future. These amendments will allow for more flexibility in the time and areas fished which may result in the development of renewed interest in flatfish fisheries. Thus the AFSC desires an independent peer review of these stocks to assess the quality of the assessments and to ensure that the Council bases its decisions on the best available information and analysis.

There are currently 21 flatfish species managed under the Bering Sea and Aleutian Islands Fishery Management Plan, of which six species have age-structured stock assessments. In the Gulf of Alaska there are ten managed species, of which four have age-structured assessments. These assessments have never undergone a CIE review. While the current Council review process is in place, there remains a compelling need for an independent peer review of the Alaska flatfish assessments.

The CIE assessment review requires a total four reviewers who are thoroughly familiar with various subject areas involved in stock assessment, including population dynamics, separable age-structured models, harvest strategies, survey methodology, and the AD Model Builder programming language. They should also have experience conducting stock assessments for fisheries management. The reviewers will travel to Seattle, Washington, to discuss the stock assessments with the flatfish assessment authors and other scientists at the AFSC involved in flatfish stock assessment. Overview presentations will be made on the topics listed below and assessment authors will be available for questions from reviewers. Specific details on a few (3-4) selected stock assessments will be presented to assist in the review.

Three of the reviewers shall generate individual reports. The fourth reviewer shall generate a Summary Report that compiles the points made by the three individual reviewers into one succinct document. The individual reports shall be appended to the Summary Report, thereby providing the complete detailed information from the individual reviewers. All reports shall address the following points.

- The strengths and weaknesses of the modeling efforts for Bering Sea/Aleutian Islands and Gulf of Alaska flatfish assessments and harvest recommendations. Specifically, the review shall evaluate:
    - The analysts' use of fishery dependent and fishery independent data sources in the assessments;
    - Gaps or inconsistencies in the population dynamics modeling methodology or logic;
    - If uncertainties in assessment model results are appropriately applied to management advice; and
    - Whether the assessments provide the best available science.

Additionally, the review shall (to the extent practical) evaluate the strengths and weaknesses of:

- The effort to incorporate ecosystem indicators and shifts in states of nature in the assessments. These include modeling survey catchability with annual bottom water temperature and using the Ocean Surface Current Simulation Models (OSCURS) to define putative oceanic productivity regimes.

- The harvest control rules adopted for Bering Sea yellowfin sole and northern rock sole (where a stock-recruitment model and F_{MSY} quantities are estimated) compared to other flatfish stocks where proxy values are used. Specifically, comments on the trade-offs between the different approaches are required.

The AFSC will provide copies of stock assessment documents, survey reports, and other pertinent literature on a web site.

Specific

1. Read and become familiar with the relevant documents provided to the reviewers.. The ten age-structured assessments that encompass the focus of the review are presented in 512 pages (of which approximately 1/3 is text).
2. Discuss the stock assessment with the lead assessment scientist and survey scientists in Seattle, Washington, from June 11 to June 14, 2007.
3. No later than June 29, 2007, submit a written report of findings, analysis, and conclusions. More details on the report outline and organization are provided in Annex I. The report shall be sent via e-mail to Dr. David Die (ddie@rsmas.miami.edu) and to Mr. Manoj Shivlani (mshivlani@rsmas.miami.edu).

Annex 1: Contents of Reviewer Reports

The following requirements refer to all reports, both the individual reports and the Summary Report.

1. All reports shall be prefaced with an executive summary of findings and/or recommendations.

2. The main body of all reports shall consist of a background, description of review activities, summary of findings, conclusions/recommendations, and references.

3. The reports shall also include as separate appendices the bibliography of all materials provided and any additional papers cited, along with 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/cimas/Report_Standard_Format.html

Review by Dr. Graham Pilling

Peer Review of Alaska flatfish stock assessments, June 11-15 2007

for

Centre for Independent Experts (NTVI)

June 2007

Cefas Contract

C2958 COMMERCIAL IN CONFIDENCE

Executive Summary

Flatfish fisheries in Alaska are valuable and are likely to receive increased interest from commercial fishers in the next decade. Most flatfish populations are currently at stable or high levels of abundance. The North Pacific Fisheries Management Council (Council) is likely to pass amendments to rationalize (establish quota shares) groundfish fisheries in the near future. These amendments will allow for more flexibility in the time and areas fished which may result in the development of renewed interest in flatfish fisheries. Thus the AFSC desired an independent peer review of these stocks to assess the quality of the assessments and to determine whether the Council is being provided with the best available information and analysis.

There are currently 21 flatfish species managed under the Bering Sea and Aleutian Islands Fishery Management Plan, of which six species have age-structured stock assessments. In the Gulf of Alaska there are ten managed species, of which four have age-structured assessments. These assessments have never undergone a CIE review. While the current Council review process is in place, there remains a compelling need for an independent peer review of the Alaska flatfish assessments.

The review was convened during June 11-14th 2007 in Seattle, Washington, at the NMFS Alaska Fisheries Science Centre. The data available and the stock assessments and forecast procedures, management protocols and research and development activities for thirteen different stocks (many grouped into 'deep water' and 'shallow water' guilds in the Gulf of Alaska) across the Bering Sea/Aleutian Islands and Gulf of Alaska were presented to the Panel, and the issues considered through open discussion. The review panel examined these issues with reference to their specific Terms of Reference (Appendix 1).

This reviewer finds that the stock assessment scientists should be commended in developing and pursuing assessments for these stocks, and I would like to thank the stock assessment team for their openness and responsiveness to questions during the review meeting, and the clarity of their reporting. The assessment approaches applied to the flatfish stock data appear to provide the best available science (ToR 1d). Specific recommendations from this reviewer resulting from the review are presented here, by Terms of Reference, and the reader should refer to the main text for full context.

ToR 1a

Current assessments rely on the abundance estimates developed from standardised fishery independent surveys, as the best information available. A particular cause for concern was the potential cessation of these surveys for budgetary reasons. Recommendations therefore focused on ways to mitigate this. The issue of whether catch information fully incorporated potential catches outside the U.S. fishing zone was also considered.

Recommendation 1. Examine the commercial and observer time series of catches as an alternative indicator of abundance for use in the future when survey time series may become limited.

Recommendation 2. Determine the robustness of the outputs of flatfish assessment approaches to uncertainty in commercial and observer catch data, using Management Strategy Evaluation. For example, identify the level of data uncertainty at which assessments become biased or imprecise, and whether this level is felt representative of the current data available. If so, how can the data collection be improved (e.g. how complete does observer coverage need to be)?

Recommendation 3. Some uncertainty exists over the level of catches of flatfish in neighbouring waters. Current assumptions (e.g. Greenland turbot) should be evaluated, and investigations pursued where existing knowledge suggests the level of catches in neighbouring waters may be important. Where necessary, collaboration with neighbouring countries during assessments is recommended, although it is realised this is easier said than done! The importance of collaboration can be examined using Management Strategy Evaluation. Differences are likely most important under appreciable exploitation levels and as Kell *et al*. (2004) identified for ICES assessment approaches, considerable bias in the perceived state of a stock can result, particularly where the management regimes in two areas are highly different.

Recommendation 4. Zones within the U.S. management area are considered discrete, with no migration between them. The limited tag returns provide little information on the potential confounding effects of movement between the Bering Sea, Aleutian Islands and Gulf of Alaska on the catch data. Again, Management Strategy Evaluation could be used to examine the significance of different migration hypotheses on the effectiveness of management.

Tor 1b

Current assessment models are developed in AD Model Builder, and some comments were raised on the ease with which this package can fit a large number of parameters. Current approaches to ensure no confounding effects when fitting parameters were adequate, although consideration of approaches for future years where the number of parameters fitted increased was warranted. Particular recommendations focussed on assumptions made on individual parameters or sub-models, and the use of retrospective analyses and projections, as well as comments on individual assessments.

Recommendation 5. Routine performance of retrospective analyses of the stock assessments is recommended.

Recommendation 6. Given the influence of the form of the stock-recruitment relationship used within assessments, sensitivity runs or preferably Management Strategy Evaluation should be performed to look at the robustness of current assessments and management to this uncertainty (see also discussions under ToR 3). Furthermore, there is the potential to model temperature impacts on recruitment (as examined under ToR2) through effects on the stock-recruitment relationship, which can be examined further using Management Strategy Evaluation (e.g. Kell *et al*., 2005).

Recommendation 7. The value of natural mortality used is constant over sexes and ages, which is unlikely to be biologically realistic. Age-specific natural mortality estimates can be developed using the MSVPA estimates of predation mortality at age, or Lorenzen's formula. The impact of this on the assessment results, stock recruitment relationship, and current tier 5 calculations should be examined.

Recommendation 8. Assessment models use mean weight-at-age whereas growth is known to be sex-specific. This may bias results where selectivity is sex-preferential. Examine the implications of using mean weight-at-age within current assessments using Management Strategy Evaluation where the operating model is sex-specific. This will require appropriate conditioning of the operating model to produce historically observed data in the perceived model.

Recommendation 9. Continue to progress the development of multispecies projection models to allow more explicit consideration of the interactions between species and fisheries to be considered in management advice. This is of particular importance given that halibut bycatch levels often limit the flatfish fishery. If sex-disaggregated assessment models are found to be beneficial, sex-specific projection models should be considered.

Recommendation 10. The survey data for Greenland turbot covers the period after high exploitation. Catch data covers the high exploitation period (although a number of assumptions are made), while later catches will be influenced by management restrictions and hence will not represent actual abundance. It is difficult to identify the impact of these issues on assessment results (including the stock-recruitment data, ABC and OFL calculations). The impact of the assumptions made and settings selected when fitting the SS2 model to the time periods of catch and survey data available for Greenland turbot should be examined. If this has already been performed, it would be helpful to include a short summary of these within the assessment report for future reviews.

Recommendation 11. The 2005 analysis of Gulf of Alaska rex sole was based on intermittent age information up to 1996. Given notable changes in length-at-age have been seen in historical data, available age information for years post 1996 should be processed to ensure any trends in length-at-age are incorporated within the assessment.

Recommendation 12. For Bering Sea and Aleutian Islands yellowfin sole, some uncertainty exists over the accuracy of early ages based upon whole otolith readings. Examine the impact of relaxing model fitting constraints on early years of the catch age structure time series for yellowfin sole on stock assessment results.

Tor 1c

Uncertainty is generally reported at the assessment stage, and is implicit within the management tier system used. While a discussion of areas where uncertainty can be included within results is developed, it is realised that the assessment process occurs within a wider management context that has strong implications for the inclusion and reporting of uncertainty.

Recommendation 13. While the assessment and projection conform to National Standard Guidelines, consider providing a better indication of uncertainty within the assessment and projection model, for example inclusion of uncertainty in current year numbers-at-age and use of the estimated stock-recruitment relationship with accompanying uncertainty. However, note that this process should be performed in full dialogue with the managers that will use the assessment results!

ToR 2

The work performed to incorporate ecosystem indicators and shifts in the state of nature in assessments is appropriate. The work examining the effect of temperature on survey catchability is developing well, and some suggestions are made of areas that can be considered within the modelling. The larval dispersal model represents good science. The next stage will be to examine how these results can be developed to understand the key developmental stages between larvae and recruitment to the survey or fishery, in order to tie the study with the assessment. The multispecies work being undertaken has potential to provide useful results for consideration alongside single species assessments.

Recommendation 14. The work modelling the impact of temperature on survey catchability may benefit from developing an underlying metabolic hypothesis to drive model development. Examine the use of metabolic theory to develop testable hypotheses to better explain the impact of temperature on survey catchability.

Recommendation 15. Consider examining the impact of temperatures on fleet distributions using available VMS data. When linked to the flatfish abundance/temperature models, the fleet distribution model may help assess the potential overlap of fleet and flatfish stock distributions. This may give insights on the future level of bycatch of flatfish stocks as a result of climate change.

Recommendation 16. Develop research programmes to identify and understand the key developmental stages for flatfish and their associated processes, and/or identify those stages key to assessment and management through Management Strategy Evaluation in order to focus this research. This should aim to link the larval dispersal model results to recruitment levels.

Recommendation 17. Examine routes to progress multispecies advice from the qualitative advice currently given (which it must be noted is already a step beyond many fisheries organisations!), to more integrated quantitative advice. MSVPA and multispecies projection models tie in directly with current assessment methods and prescribed management approaches, while Ecopath with Ecosim studies provide a wider view of the multispecies ecosystem in the region that may be appropriate for future management requirements. Both general approaches are, in my view, worthy of continued investigation.

ToR 3

The current tier system for management is an excellent idea in principle. It would benefit from a clear documentation of the reasons behind assignment of individual stocks to a tier, for future reference. Comments are made on the tier system in general, with particular focus on the simulation testing of the tier harvest rules to ensure they work as expected. Ultimately, designation of decision rules of this kind requires the definition of acceptable risk, which is a role for managers, fishermen and other stakeholders to derive. Fisheries assessment scientists can then test the performance of proposals based on the available knowledge of the system. Further recommendations are developed for individual stocks.

Recommendation 18. Use Management Strategy Evaluation to ensure that the tier decision system is working as expected with respect to the robustness and precautionary nature of advice in given data situations. Also examine the performance of the harmonic mean estimate of F_{msy} under current and alternative assumptions of model uncertainty.

Recommendation 19. Examine the implications of uncertainty in natural mortality on the performance of tier 5 management rules. Is uncertainty in the natural mortality estimate critical?

Recommendation 20. Based on the collection of new age information, analyses of alternative contributing factors (e.g. the use of excluders in gear) etc., examine the implications of alternative selectivity and maturity ogives on assessment results for Gulf of Alaska rex sole.

Specifically for the Bering Sea and Aleutian Islands yellowfin sole and Northern rock sole, this reviewer finds the stock recruitment relationships developed to be as good a fit as possible to the data available, and given the selected functional form of that relationship. Therefore I see no reason not to move these species to tier 1.

Background

Flatfish fisheries in Alaska are valuable and are likely to receive increased interest from commercial fishers in the next decade. Most flatfish populations are currently at stable or high levels of abundance. The North Pacific Fisheries Management Council (Council) is likely to pass amendments to rationalize (establish quota shares) groundfish fisheries in the near future. These amendments will allow for more flexibility in the time and areas fished which may result in the development of renewed interest in flatfish fisheries. Thus the AFSC desired an independent peer review of these stocks to assess the quality of the assessments and to determine whether the Council is being provided with the best available information and analysis.

There are currently 21 flatfish species managed under the Bering Sea and Aleutian Islands Fishery Management Plan, of which six species have age-structured stock assessments. In the Gulf of Alaska there are ten managed species, of which four have age-structured assessments. These assessments have never undergone a CIE review. While the current Council review process is in place, there remains a compelling need for an independent peer review of the Alaska flatfish assessments.

This document represents the individual CIE Reviewer Report on the results of the review panel deliberations on the Alaska flatfish stock assessments, at the request of the Center for Independent Experts (see Appendix 1). The author was provided with the stock assessment reports prior to the meeting (see bibliography), assimilated the presentations and participated fully in the discussions at the review meeting, which formed a key part of the review panel process.

Description of review activities

The review was undertaken by Dr Graham Pilling at Cefas (Lowestoft, UK) and during the Review Panel held in Seattle, Washington, at the NMFS Alaska Fisheries Science Centre. The Flatfish Review Panel was convened during June 11-14th 2007. The panel membership is listed in Appendix 2.

The documentation (see bibliography) was reviewed at Cefas, prior to travel. Dr Pilling actively participated in the panel meeting in Seattle. This report to CIE was completed on return to Cefas.

The data available and the stock assessments and forecast procedures, management protocols and research and development activities for thirteen different stocks (many grouped into ‘deep water’ and ‘shallow water’ guilds in the Gulf of Alaska) across the Bering Sea/Aleutian Islands and Gulf of Alaska were presented to the Panel, and the issues considered through open discussion. The review panel examined these issues with reference to their specific Terms of Reference (Appendix 1).

Summary of findings

The four-day meeting of the Peer Review Panel of Alaska Flatfish Stock Assessments covered the data collection, assessment, management and research and development activities for thirteen different stocks across the Bering Sea/Aleutian Islands and Gulf of Alaska. The stock assessment scientists should be commended in developing and pursuing assessments for these stocks, which while currently of relatively minor importance in terms of directed fisheries in the region, have considerable potential importance from an ecosystem perspective and as future fisheries. This reviewer would like to thank the stock assessment team for their openness and responsiveness to questions during the review meeting, and the clarity of their reporting.

Directed fisheries for these flatfish stocks can be limited by the bycatch levels of species such as halibut (e.g. Barents Sea rock sole, flathead sole) and red king crab (e.g. Barents Sea yellowfin sole). As a result, the TAC level assigned to stocks is often not fully taken.

It is noted that the future of surveys that are used to give flatfish abundance information (although the surveys are not targeted at these species) is uncertain. This is a concern, since they provide an excellent source of information on the biology and distribution of these species, and form the basis of a number of studies. Recommendations are made within the sections below on potential ways to mitigate this problem, but the ultimate recommendation would be to maintain the time series of these surveys if at all possible.

The findings of this reviewer are reported within relevant sections, addressing the three main areas of the Review Terms of Reference (Appendix 1). Numbered recommendations (in bold) refer to the correspondingly numbered items within the conclusions and recommendations section of this report.

1. The strengths and weaknesses of the modeling efforts for Bering Sea/Aleutian Islands and Gulf of Alaska flatfish assessments and harvest recommendations. Specifically, the review shall evaluate: a) the analysts' use of fishery dependent and fishery independent data sources in the assessments; b) gaps or inconsistencies in the population dynamics modeling methodology or logic; c) if uncertainties in assessment model results are appropriately applied to management advice; and d) whether the assessments provide the best available science.

a) Analysts' use of fishery dependent and fishery independent data sources in assessments

The use of fishery independent data sources is discussed first, and observations on the fishery dependent data second.

Currently, fishery independent data sources are used to estimate trends in flatfish abundance. While not targeting flatfish specifically, these surveys are perceived to provide good standardised indices for relative biomass estimation of these species, a reflection of the limited targeting of flatfish in commercial fisheries, and the impacts of bycatch species on the commercial fishery.

It was good to see that investigations on the fishery independent survey continued. In particular, the potential herding effect of warps and the implications for the catchability of flatfish were being examined, since this effect could strongly influence the abundance estimates upon which the stock assessments rely. The use of cameras within these studies was implied, and these provide direct observations of the behaviour of flatfish species on encountering the gear, including herding and escapement patterns. Camera observations could also provide additional information on the implications of water temperature on the catchability of flatfish in surveys, which was reported during the review. The studies underway to examine this effect are discussed further under ToR 2.

A cause for concern is the potential cessation of the fishery independent survey for budgetary reasons. As a result, there is a need to look at the time series of information from the commercial fishery, and in particular the time series of data from the commercial observer programme, as an alternative index of abundance. Observer coverage is not necessarily at random (although the aim is to achieve a random sample of trawls from those vessels and trips observed). Efforts to standardise for gear and vessels effects will be needed, where sufficient information is available. **See recommendation 1**.

The information available on commercial catches is considered to be good. Misreporting is likely minimal for flatfish, as fishing opportunities are generally not limited by the catch of these species (rather by the bycatch). Discarding is felt reasonably recorded (and is included within assessments) due to observer coverage. While there is potential for bias in catch data, it is recognised, if difficult to incorporate. Currently, the emphasis placed upon the survey information means this is less of an issue. With the potential cessation of the survey, this may become a greater issue and the robustness of assessment models to this bias should be determined. **See recommendation 2**.

The catch information is constrained to the US zone, and catch information is not directly available from the bordering Russian and Canadian zones. Given the low exploitation levels within the flatfish fishery, tagging returns will be sporadic and likely to provide limited information. The current opportunistic use of existing tags, rather than a specific tagging programme, is therefore appropriate. Anecdotal evidence suggests that the catches of flatfish in neighbouring waters are minimal, based on the limited abundance levels in US waters neighbouring Canada, and the small seabed shelf area in the Russian zone. It was noted that the survey might not cover the whole stock of Greenland turbot, as this species is found on the slope, which extends into the Russian zone (current assessments using SS2 assume that the U.S. waters slope-trawl survey abundances represent 75% of the total stock). Limited tagging information available suggests that some movement may occur between the US and Russian zones. Closer to home, there is also a lack of knowledge on the movement of fish within zones of the U.S. management area. Results from the Ecopath model suggest that the species composition between the Gulf of Alaska and Bering Sea are quite different, which may suggest minimal migration. There is a need to identify the degree to which species distribution may overlap between management zones, particularly where the centroid of abundance may be affected by changes in temperature regime, or identify the impact of current assumptions upon the performance of the assessment. **See recommendations 3 and 4**.

b) Gaps or inconsistencies in the population dynamics modelling methodology or logic

The AD Model Builder approach used for flatfish has been developed as a ‘generic’ model that can be applied to all species, given available information. For some species, simpler or alternative approaches such as SS2 have been used. The use of appropriate models for different species with different available information is recommended, since this allows the best potential stock assessment to be developed predicated upon the information available. The potential problem with developing a truly generic model is the need to build in such levels of flexibility that the model becomes too complex to be easily used.

AD Model Builder offers the potential to fit a large number of parameters within a model. However, with a large number of parameters care needs to be taken to ensure no confounding effects are occurring between those parameters. The inverse Hessian matrix and MCMC evaluation of the posterior are both routinely examined for confounding between parameters or poor estimation. As additional years of data are collected, and further parameters need to be estimated, this process will become more and more difficult, and some consideration to how this will be dealt with in the future should be given.

Although comparisons are made between the results of the current years’ assessment and that run in the previous assessment, no formal retrospective analyses are performed on the assessments. These are invaluable to identify persistent assessment biases and uncertainties within the data or assessment performed. **See recommendation 5**.

The Ricker stock-recruitment relationship has been used throughout the assessments, based upon the criteria of best fit to the available data. Within ICES, the Beverton and Holt relationship has generally been applied, due to the lack of biological evidence for cannibalism or stock limitation at higher population sizes in those stocks. However, the understanding is that the Ricker relationship is likely to be more precautionary, at least at the high stock sizes currently estimated. The impact of stock-recruitment relationship model uncertainty on the performance of management should be examined through simulation testing, and care must be taken where stock sizes decrease. **See recommendation 6**.

Where fishing mortality is low, there is potentially little contrast within the available data from which the model can develop an accurate fit. This does raise some uncertainty over whether the level of fishing/total mortality is being estimated well. The model does fit the age information very well, adding credence to the model results. The lack of returns from tagging programmes, analysis of catch curves (as performed by reviewer Dr Paul Medley following the review panel) and general age structure of the catches suggests that fishing mortality is indeed low.

A source of uncertainty within the flatfish models was the value selected for natural mortality, which was generally constant between sexes and ages. Differential natural mortality at size (rather than age) might explain some of the sexual dimorphism seen in the flatfish (e.g. Gulf of Alaska arrowtooth flounder). Furthermore, values for natural mortality at age could be developed from the values estimated within the MSVPA studies (as performed by ICES for North Sea cod, for example), or through use of Lorenzen's M equation (see Lorenzen, 1996). Although the impacts of predation mortality occur before recruitment to the fishery, they can influence the level of age 1 recruitment estimates. It also has potential impact on the calculations performed within the management tier system (see discussions under ToR 3). **See recommendation 7**.

A 'plus group' has generally been used within assessments. Given the perceived low level of fishing mortality for the flatfish species, there tends to be a build-up of individuals within this plus group (e.g. Bering Sea and Aleutian Islands yellowfin sole and Greenland turbot). Increases in length tend to be relatively small at this age or length, and the impact of using a mean weight-at-age or –at-length for the plus group (and the impact on spawning biomass) is also likely to be minor, but worth monitoring.

Most models assume length- or weight-at-age to be the mean of males and females where growth is sexually dimorphic. This may prove reasonable, but has the potential to cause a bias where selectivity is sex-preferential (e.g. towards females due to faster growth and greater size). The degree of bias cannot be easily identified, as it depends on the fishing pressure (which is relatively low), the assessment model used (single sex) and the management system in operation. Hence this is another appropriate study within a Management Strategy Evaluation framework. **See recommendation 8**.

Currently, stock projections are performed within a separate, specifically developed generic projection model. The assessment and projection models are consistent in their basis, in general using the same equations. The exception on the latter (but not the former) is the Greenland turbot assessment, where the projection model uses age-specific future selectivity, whereas the assessment model is based upon size-selectivity (J. Ianelli, pers. comm.). Although it might be 'neater' to extend the AD Model Builder assessment model to perform projections, this is not absolutely necessary. In turn, the current generic model aims to allow the linking of bycatch information (which represents the current extent of the fisheries management system's consideration of the multispecies nature of fisheries) with single species assessments. The progression of this multispecies projection model is recommended. If sex-specific assessment models are developed following analyses discussed above, the benefits of sex-specific projections might be considered. **See recommendation 9**.

The survey data for Greenland turbot covers the period after high exploitation. Catch data covers the high exploitation period (although a number of assumptions are made to generate this), while later catch levels will be influenced by management constraints and hence will not represent actual abundance. It is difficult to identify the impact of these issues on assessment results (including the stock-recruitment data, ABC and OFL calculations). **See recommendation 10**.

The 2005 analysis of Gulf of Alaska rex sole was based on intermittent age information up to 1996, although length information was available from every other year up to 2005. Given that some notable changes have occurred in length-at-age between 1984 and 1990, it would be useful to obtain additional age information for more recent years to ensure that no biases occur in the assessment as a result of any unidentified trends occurring. **See recommendation 11**.

Some concerns were raised on the ageing of Bering Sea and Aleutian Islands yellowfin sole in the early years of the time series. Specifically, whole otoliths were used to age this species, which has the potential to underestimate the age composition of catches. This had been examined, and the younger age structure in the data at that time was commensurate with the high fishing mortality being experienced. However, there is the potential to relax the model constraints to the fit on these older data, to examine the effect on model results and the stock-recruitment relationship. **See recommendation 12**.

c) If uncertainties in assessment model results are appropriately applied to management advice

Uncertainty within the process is generally reported at the assessment stage. Alternative ABC and OFL levels may be provided by assessment authors based upon different data sources etc. (e.g. Gulf of Alaska rex sole). Furthermore, the level of uncertainty in the assessment defines the management tier into which the stock is placed and hence the harvest rule used. In this fashion, uncertainty in assessment model results is appropriately applied to management advice (although see response to ToR3 and recommendation 18). As in most stock assessments around the world, however, formal representation of uncertainty in assessment outputs could be increased. The ABC provided to managers is a single number, with no confidence intervals, which is generally how managers like it! Presenting fuller measures of uncertainty must trade off against the level of uncertainty in results that managers will 'stand', which must be clarified through dialogue. Further consideration of this is outside the boundaries of this review!

The projection model used following the assessment provides an upper bound for the harvest rule. While projections conform to National Standard Guidelines, considerations of uncertainty are limited to recruitment variation, which impacts the projections a number of years after the initial projection year (dependent on the age at which fish recruit to the fishery). A better consideration of uncertainty could be included through the inclusion of uncertainty in current year numbers-at-age, for example. This could have a direct impact on the shorter-term management benchmarks (e.g. overfished and overfishing criteria). Issues with tier 1 species focus on the use of the harmonic mean F_{msy}, and are discussed under ToR3. For tier 3 species, where a reliable stock recruitment relationship is available, this can be used with its associated uncertainty to bootstrap future recruitments in projections (although it is noted that using mean recruitment levels in projections is felt conservative under current high stock conditions). Note that addition of uncertainty in projections will again require an appropriate approach for presenting those results to managers to be developed. **See recommendation 13**.

d) Whether the assessments provide the best available science

As stated earlier, the work performed is impressive. The assessment approaches applied to the flatfish stock data appear to provide the best available science, and accompanying evidence that the level of fishing mortality is indeed low (catch curve analysis of Paul Medley, age structure in general, low rate of tag returns) provides further support to the results. It would be useful if alternative assessment methods such as catch curves (or VPA if fishing mortality increases) were applied at regular intervals to further support assessment model outputs and examine model uncertainty.

The potential for confounded parameter estimates where many different parameters are being estimated has been examined through the use of appropriate diagnostics. The use of retrospective analyses has been recommended (see recommendation 5).

The use of Management Strategy Evaluation has been appropriate to test particular uncertainties within the system, and recommendations to expand this work have been suggested (see many recommendations in this report). Assessment scientists should prioritise the potential simulation tests given available time.

2. Review the efforts to incorporate ecosystem indicators and shifts in states of nature in the assessments. These include modeling survey catchability with annual bottom water temperature and using the Ocean Surface Current Simulation Models (OSCURS) to define putative oceanic productivity regimes.

Given evidence of the effect of temperature on survey results, the pursuit of methods to incorporate the impact of survey catchability is worthwhile. Appropriate alternative models are being examined and the work looks to be developing very well, and being appropriately disseminated at international meetings. The work may benefit from developing an underlying metabolic hypothesis to drive model development. Anecdotal evidence suggests that temperature affects the movement levels of flatfish in the presence of the trawl gear. Brown *et al*. (2004) present a number of formulations linking the metabolism of organisms (linked to temperature, body mass etc.) to the ecology of populations, communities, and ecosystems. This underlying metabolic theory could be used to develop testable hypothesis models of the impact of temperature on catchability, linked to fish movement and population density. **See recommendation 14**.

Modelling of the effect of temperature on flatfish distribution has been undertaken using survey data. A time series of fishery VMS data is also available which can be used to examine any indirect impacts of temperature on the distribution of fishing fleets in a similar way. In theory fleet and stock distribution will overlap where the stock is being targeted. Many flatfish species are bycatch, however. The degree of overlap between fleet and stock abundance under scenarios of future temperature change may give some idea of the relative levels of bycatch and discarding of flatfish under these different climate hypotheses. **See recommendation 15**.

The larval dispersal model, driven by the Ocean Surface Current Simulation Models, is an interesting scientific study. The distribution patterns of eggs will be determined by their buoyancy, and hence the interplay between current (deeper water) and wind (shallow water) directions. The use of a larval survey to ground-truth the distribution results of the model would be highly useful to confirm the assumptions made. However, the issue will be how to tie these model results into subsequent fisheries stock assessments. This will require a better understanding of post-settlement processes, developmental rates within different temperature regimes, the potential for density dependent growth and mortality, and the survival of juveniles and young adults prior to entry to the fishery (see Bailey *et al*., 2005). Larval rearing studies (e.g. Fox *et al*., 2003) may be used to identify many of the potential impacts (and build upon the findings of Nichol and Acuna, 2001), while temperature might also influence the duration of spawning (e.g. Rijnsdrop and Witthames, 2005). Furthermore, improving the understanding of the relative impact of key stages of development on recruitment (e.g. Nash and Dickie-Collas, 2005) is a worthwhile pursuit. The potential impact of future environmental change on key developmental stages and stock status, assessment and management can also be examined through Management Strategy Evaluation. This could identify the developmental stages with the greatest potential influence on future recruitment and assessment performance, and hence identify areas to focus further research. **See recommendation 16**.

The Ecopath with Ecosim studies that have been performed appear worthwhile, and the precision of the model should improve with the continued collection of stomach contents data from observer programmes. The results demonstrate the importance of the flatfish complex as a whole, although the role of individual species appears more muted. At present, the results provide qualitative grounds for multispecies stock assessment advice for particular flatfish groups. The development of the MSVPA and multispecies projection models (the latter incorporating interactions with key species such as halibut) provides results of direct relevance to current fisheries management. **See recommendation 17**.

3. Review the harvest control rules adopted for Bering Sea yellowfin sole and northern rock sole (where a stock-recruitment model and FMSY quantities are estimated) compared to other flatfish stocks where proxy values are used. Specifically, comments on the trade-offs between the different approaches are required.

The current tier system for management provides alternative approaches to define overfishing levels (OFL) for species and stocks consistent with the Magnuson-Stevens Fishery Conservation and Management Act (Magnuson-Stevens Act). The tier system is designed to cope with stocks having different levels of reliable information for determining OFLs.

The tier system is an excellent idea in principle, since the theory behind it gives an incentive for fishermen to provide more and better data to assessment scientists in order to allow a greater level of sustainable catch based upon the better estimation of stock status that should result. Where data are lacking, decisions are expected to be more precautionary than in cases where more data are available.

The SSC defines the position of a given species within the tier system. The decisions made on tier positions are documented in SSC reports, although the reasoning behind those decisions may not be explicitly detailed, and this could be improved.

The actual performance of the tier system should be evaluated using Management Strategy Evaluation, to ensure the harvest rules at each tier are working as anticipated. A stock with good information could be modeled, and the performance of management at each tier examined to ensure the tier rules work as expected in terms of robustness to uncertainty (i.e. lower tiers are more precautionary than higher tiers). Performance criteria to assess the harvest rules could include:

- TAC level applied compared to the optimal performance TAC;
- Perceived SSB compared to the actual SSB level;
- Perceived and actual recruitment levels;
- Inter-annual variability in catches;

- Proportion of years during the given period that the stock meets/exceeds an overfished criterion;

which can then be looked at in the short, medium and long term. This approach can also be used to test whether the use of the harmonic mean estimate of F_{msy} provides a sufficient buffer between the OFL and ABC levels, which may be strongly influenced by the level of uncertainty within the assessment model. It can also examine whether the use of the Ricker stock recruitment relationship and associated reference points is more conservative - as anticipated - compared to use of the Beverton and Holt model. **See recommendation 18**.

Designation of decision rules for the different tiers ultimately requires a formal risk analysis. Definition of acceptable risk is not the role of fisheries assessment scientists, but of managers, fishermen, and other stakeholders to derive. Fisheries assessment scientists can then test the performance of these decision rules based on knowledge of the system in which the framework operates.

On the current system, the general harvest rules at each tier appear sensible. However, it should be noted that the natural mortality parameter is seldom estimated with any real reliability. Techniques based upon the oldest age in the population (which is obviously influenced by the level of fishing pressure experienced), empirical approaches or meta-analyses are generally accepted, but alternative criteria could also be tested for the tier 5 criterion. For example, comparison of the length at capture (Lc_{50}) and length at maturity (Lm_{50}) - a spawn once before capture policy - could be a good signal, despite the high levels of effort that could result where $Lc_{50} >> Lm_{50}$ (e.g. Gulf of Alaska rex sole).

For tier 5 species (e.g. Gulf of Alaska shallow water flatfish) where natural mortality is less certain, the assumption is taken that natural mortality is 0.2. However, in the other flatfish species where natural mortality is estimated within the assessment model, natural mortality is estimated to be under this value. The implications of this assumption should be examined versus the implications of dropping such species into the tier 6 category (given the potential lack of a ‘reliable’ natural mortality estimate). **See recommendation 19.**

Specifically for the Bering Sea and Aleutian Islands yellowfin sole and Northern rock sole, this reviewer finds the stock recruitment relationships developed to be as good a fit as possible to the data available, given the selected functional form of that relationship. Therefore I see no reason not to move these species to tier 1.

For the Gulf of Alaska rex sole assessment, concern was raised over the high F that was allowable due to the relative lengths of capture and maturity. It was felt that at such high levels of mortality, which imply a targeted fishery, the selectivity of the commercial gear would change to exploit younger fish, and lead to overexploitation at that high F. Care was therefore warranted. Collecting new age and maturity information might reduce uncertainty over the estimates of maturity and selectivity at age (see recommendation 12). Further issues, including the use of halibut excluders and seasonal differences between surveys and the commercial fishery were implicated, but it is not known whether these have been investigated. The implications of changes in selectivity could be tested to inform management. Using implied effort levels calculated where the lengths at maturity and capture were equal would currently be highly precautionary. Alternative relative parameter values could be based upon those for other flatfish species which are the subject of more targeted fisheries (e.g. Greenland turbot) where exploitation levels are higher and the length at capture may have changed as a result, or developed by asking fishermen about the fish sizes that would be most profitable commercially. Obviously, care must be taken with the latter, since market drivers (and price-at-size dynamics) change when species are targeted. **See recommendation 20**.

4. General comments

The work presented on otolith increment validation was very good. There is the potential to use oxygen isotope ratios (and the resulting temperature signal) using micromill techniques for validation at younger ages where sufficient material can be gathered and resolution is sufficient not to integrate over too much time/material (see Hoie and Folkvord, 2006; Hoie *et al*., 2004).

5. Conclusions/Recommendations

The stock assessment scientists should be commended in developing and pursuing assessments for these stocks, and I would like to thank the stock assessment team for their openness and responsiveness to questions during the review meeting, and the clarity of their reporting.

The assessment approaches applied to the flatfish stock data appear to provide the best available science. Specific recommendations from this reviewer resulting from this review are presented here, and the reader should refer to the main text for full context.

ToR 1a

Recommendation 1. Examine the commercial and observer time series of catches as an alternative indicator of abundance for use in the future when survey time series may become limited.

Recommendation 2. Determine the robustness of the outputs of flatfish assessment approaches to uncertainty in commercial and observer catch data, using Management Strategy Evaluation. For example, identify the level of data uncertainty at which assessments become biased or imprecise, and whether this level is felt representative of the current data available. If so, how can the data collection be improved (e.g. how complete does observer coverage need to be)?

Recommendation 3. Some uncertainty exists over the level of catches of flatfish in neighbouring waters. Current assumptions (e.g. Greenland turbot) should be evaluated, and investigations pursued where existing knowledge suggests the level of catches in neighbouring waters may be important. Where necessary, collaboration with neighbouring countries during assessments is recommended, although it is realised this is easier said than done! The importance of collaboration can be examined using Management Strategy Evaluation. Differences are likely most important under appreciable exploitation levels and as Kell *et al*. (2004) identified for ICES assessment approaches, considerable bias in the perceived state of a stock can result, particularly where the management regimes in two areas are highly different.

Recommendation 4. Zones within the U.S. management area are considered discrete, with no migration between them. The limited tag returns provide little information on the potential confounding effects of movement between the Bering Sea, Aleutian Islands and Gulf of Alaska on the catch data. Again, Management Strategy Evaluation could be used to examine the significance of different migration hypotheses on the effectiveness of management.

Tor 1b

Recommendation 5. Routine performance of retrospective analyses of the stock assessments is recommended.

Recommendation 6. Given the influence of the form of stock-recruitment relationship used within assessments, sensitivity runs or preferably Management Strategy Evaluation should be performed to look at the robustness of current assessments and management to this uncertainty (see also discussions under ToR 3). Furthermore, there is the potential to model temperature impacts on recruitment (as examined under TOR2) through effects on the stock-recruitment relationship, which can be examined further using Management Strategy Evaluation (e.g. Kell *et al*., 2005).

Recommendation 7. The value of natural mortality used is constant over sexes and ages, which is unlikely to be biologically realistic. Age-specific natural mortality estimates can be developed using the MSVPA estimates of predation mortality at age, or Lorenzen's formula, to develop biologically more realistic natural mortality-at-age estimates. The impact of this on the assessment results, stock recruitment relationship and current tier 5 calculations should be examined.

Recommendation 8. Assessment models use mean weight-at-age whereas growth is known to be sex-specific. This may bias results where selectivity is sex-preferential. Examine the implications of using mean weight-at-age within current assessments using Management Strategy Evaluation where the operating model is sex-specific. This will require appropriate conditioning of the operating model to produce historically observed data in the perceived model.

Recommendation 9. Continue to progress the development of multispecies projection models to allow more explicit consideration of the interactions between species and fisheries to be considered in management advice. This is of particular importance given that halibut bycatch levels often limit the flatfish fishery. If sex-disaggregated assessment models are found to be beneficial, sex-specific projection models should be considered.

Recommendation 10. The impact of the assumptions made and settings selected when fitting the SS2 model to the time periods of catch and survey data available for Greenland turbot should be examined. If this has already been performed, it would be useful to include a short summary of these within the assessment report for future reviews.

Recommendation 11. The 2005 analysis of Gulf of Alaska rex sole was based on intermittent age information up to 1996. Given notable changes in length-at-age have been seen in historical data, available age information for years post 1996 should be processed to ensure any trends in length-at-age are incorporated within the assessment.

Recommendation 12. For Bering Sea and Aleutian Islands yellowfin sole, some uncertainty exists over the accuracy of early ages based upon whole otolith readings. Examine the impact of relaxing model fitting constraints on early years of the catch age structure time series for yellowfin sole on stock assessment results.

Tor 1c

Recommendation 13. While the assessment and projection conform to National Standard Guidelines, consider providing a better indication of uncertainty within the assessment and projection model, for example inclusion of uncertainty in current year numbers-at-age and use of the estimated stock-recruitment relationship with accompanying uncertainty. However, not that this process should be performed in full dialogue with the managers that will use the assessment results!

ToR 2

Recommendation 14. The work modelling the impact of temperature on survey catchability may benefit from developing an underlying metabolic hypothesis to drive model development. Examine the use of metabolic theory to develop testable hypotheses to better explain the impact of temperature on survey catchability.

Recommendation 15. Consider examining the impact of temperatures on fleet distributions using available VMS data. When linked to the flatfish abundance/temperature models, the fleet distribution model may help assess the potential overlap of fleet and flatfish stock distributions. This may give insights on the future level of bycatch of flatfish stocks as a result of climate change.

Recommendation 16. Develop research programmes to identify and understand the key developmental stages for flatfish and their associated processes, and/or identify those stages key to assessment and management through Management Strategy Evaluation in order to focus this research. This should aim to link the larval dispersal model results to recruitment levels.

Recommendation 17. Examine routes to progress multispecies advice from the qualitative advice currently given (which it must be noted is already a step beyond many fisheries organisations!), to more integrated quantitative advice. MSVPA and multispecies projection models tie in directly with current assessment methods and prescribed management approaches, while Ecopath with Ecosim studies provide a wider view of the multispecies ecosystem in the region that may be appropriate for future management requirements. Both general approaches are, in my view, worthy of continued investigation.

ToR 3

Recommendation 18. Use Management Strategy Evaluation to ensure that the tier decision system is working as expected with respect to the robustness and precautionary nature of advice in given data situations. Also examine the performance of the harmonic mean estimate of F_{msy} under current and alternative assumptions of model uncertainty.

Recommendation 19. Examine the implications of uncertainty in natural mortality on the performance of tier 5 management rules. Is uncertainty in the natural mortality estimate critical?

Recommendation 20. Based on the collection of new age information, analyses of alternative contributing factors (e.g. the use of excluders in gear) etc., examine the implications of alternative selectivity and maturity ogives on assessment results for Gulf of Alaska rex sole.

Appendix I: Bibliography

Primary documentation

Wilderbuer, T.K. and Nichol, D.G. (2006). Chapter 4. Bering Sea and Aleutian Islands Yellowfin sole. p 429-492.

Wilderbuer, T.K. and Nichol, D.G. (2006). Chapter 7. Bering Sea and Aleutian Islands Northern rock sole. p 591-652.

Turnock, B.J., Wilderbuer, T.K. and Brown, E.S. (2005). Chapter 4. Gulf of Alaska flatfish. p 317-350.

Ianelli, J.N., Wilderbuer, T.K. and Nichol, D. (2006). Chapter 5. Assessment of Greenland turbot in the Eastern Bering Sea and Aleutian Islands. p 493-540.

Turnock, B.J., Wilderbuer, T.K. and Brown, E.S. (2005). Gulf of Alaska arrowtooth flounder stock assessment. P 435-472.

Turnock, B.J. and A'mar Z.T. (2005). Gulf of Alaska rex sole stock assessment. p 399-434.

Draft report of the Scientific and Statistical Committee to the North Pacific Fishery Management Council, December 6-8, 2006.6p.

The Plan Team for the Groundfish Fisheries of the Bering Sea and Aleutian Islands. (2006). Appendix A. Stock assessment and fishery evaluation report for the groundfish resources of the Bering Sea/Aleutian Islands. North Pacific Fishery Management Council. 34p.

The Plan Team for the Groundfish Fisheries of the Gulf of Alaska. (2006). Appendix B. Stock assessment and fishery evaluation report for the groundfish resources of the Gulf of Alaska. North Pacific Fishery Management Council. 38p.

Additional references used by the reviewer

Bailey, K.M., Nakata, H. and Van der Weer, H.W. (2005). The planktonic stages of flatfishes: physical and biological interactions in transport processes. P94-119 *in* Flatfishes: biology and exploitation (ed. Gibson, R.N.). Blackwell publishing, Oxford.

Brown, J.H., Gillooly, J.F., Allen, A.P., Savage, V.M. and West, G.B. (2004). Toward a metabolic theory for ecology. *Ecology* **85**(7), 1771-1789.

Fox, C.J., Geffen, A.J., Blyth, R. and Nash, R.D.M. (2003). Temperature-dependent development rates of plaice (*Pleuronectes platessa* L.) eggs from the Irish Sea. *Journal of Plankton Research* **25**(11), 1319-1329.

Hoie, H., Andersson, C., Folkvord, A. and Karlsen, Ø. (2004). Precision and accuracy of stable isotope signals in otoliths of pen-reared cod (*Gadus morhua*) when sampled with a high-resolution micromill. *Marine Biology* **144**(6), 1039-1049.

Hoie, H. and Folkvord, A. (2006). Estimating the timing of growth rings in Atlantic cod otoliths using stable oxygen isotopes. *Journal of Fish Biology* **68**, 826–837.

Kell, L.T., Scott, R. and Hunter, E. (2004). Implications for current management advice for North Sea plaice: Part 1. Migration between the North Sea and English Channel. *Journal of Sea Research* **51**, 287-299.

Kell, L.T., Pilling, G.M. and O'Brien, C.M. (2005). The implications of climate change for the management of North Sea cod (*Gadus morhua*). *ICES Journal of Marine Science* **62**, 1483-1491.

Lorenzen, K. (1996). The relationship between body weight and natural mortality in juvenile and adult fish: a comparison of natural ecosystems and aquaculture. *Journal of Fish Biology* **49**, 627-647.

Miller, T. J., Skalski, J. R., and Ianelli, J. N. (2007). Optimizing a stratified sampling design when faced with multiple objectives – *ICES Journal of Marine Science* **64**, 97–109.

Nash, R.D.M. and Dickie-Collas, M. (2005). The influence of life history dynamics and environment on the determination of year class strength in North Sea herring (*Clupea harengus* L.). *Fisheries Oceanography* **14**(4), 279-291.

Nichol, D.G. and Acuna, E.I. (2001). Annual and batch fecundities of yellowfin sole, *Limanda aspera*, in the eastern Bering Sea. *Fishery Bulletin* **99**, 108-122.

Rijnsdrop, A.D. and Witthames, P.R. (2005). Ecology of reproduction. p68-93 *in* Flatfishes: biology and exploitation (ed. Gibson, R.N.). Blackwell publishing, Oxford.

Spencer, P.D., Wilderbuer, T.K. and Zhang, C.I. (2002). A mixed-species yield model for eastern Bering Sea shelf flatfish fisheries. *Canadian Journal of Fisheries and Aquatic Sciences* **59**, 291-302.

Appendix 1. Statement of work

CIE REQUEST

Subcontract between NTVI and the Centre for Environment, Fisheries and Aquaculture Sciences (CEFAS) (Dr. Graham Pilling)

<u>Alaska Flatfish Review Statement of Work</u>

General

The Alaska Fisheries Science Center (AFSC) requests a peer review of the Gulf of Alaska and Bering Sea/Aleutian Islands flatfish stock assessments by the Center for Independent Experts (CIE). Flatfish fisheries in Alaska are valuable and are likely to receive increased interest from commercial fishers in the next decade. Most flatfish populations are currently at stable or high levels of abundance. The North Pacific Fisheries Management Council (Council) is likely to pass amendments to rationalize (establish quota shares) groundfish fisheries in the near future. These amendments will allow for more flexibility in the time and areas fished which may result in the development of renewed interest in flatfish fisheries. Thus the AFSC desires an independent peer review of these stocks to assess the quality of the assessments and to determine whether the Council is being provided with the best available information and analysis.

There are currently 21 flatfish species managed under the Bering Sea and Aleutian Islands Fishery Management Plan, of which six species have age-structured stock assessments. In the Gulf of Alaska there are ten managed species, of which four have age-structured assessments. These assessments have never undergone a CIE review. While the current Council review process is in place, there remains a compelling need for an independent peer review of the Alaska flatfish assessments.

The CIE assessment review requires a total of four reviewers who are thoroughly familiar with various subject areas involved in stock assessment, including population dynamics, separable age-structured models, harvest strategies, survey methodology, and the AD Model Builder programming language. They should also have experience conducting stock assessments for fisheries management. The reviewers will travel to Seattle, Washington, to discuss the stock assessments with the flatfish assessment authors and other scientists at the AFSC involved in flatfish stock assessment. Overview presentations will be made on the topics listed below and assessment authors will be available for questions from reviewers. Specific details on a few (3-4) selected stock assessments will be presented to assist in the review.

Three of the reviewers shall generate individual reports. The fourth reviewer shall generate a Summary Report that compiles the points made by the three individual reviewers into one succinct document. The individual reports shall be appended to the Summary Report, thereby providing the complete detailed information from the individual reviewers. All reports shall address the strengths and weaknesses of the following points.

- Modeling efforts for Bering Sea/Aleutian Islands and Gulf of Alaska flatfish assessments and harvest recommendations. Specifically, the review shall evaluate:
    - The analysts' use of fishery dependent and fishery independent data sources in the assessments;
    - Gaps or inconsistencies in the population dynamics modeling methodology or logic;
    - If uncertainties in assessment model results are appropriately applied to management advice; and
    - Whether the assessments provide the best available science.

- The effort to incorporate ecosystem indicators and shifts in states of nature in the assessments. These include modeling survey catchability with annual bottom water temperature and using the Ocean Surface Current Simulation Models (OSCURS) to define putative oceanic productivity regimes.

- The harvest control rules adopted for Bering Sea yellowfin sole and northern rock sole (where a stock-recruitment model and F_{MSY} quantities are estimated) compared to other flatfish stocks where proxy values are used. Specifically, comments on the trade-offs between the different approaches are required.

The AFSC will provide copies of stock assessment documents, survey reports, and other pertinent literature on a web site.

Specific

The CIE shall provide four reviewers, for a maximum total of 52 work days. The three individual reviewers are approved for a maximum of 15 work days each, and the summarizer is approved for a maximum of seven work days. The three individual reviewers shall be approved for travel to the meetings at the AFSC. The summarizer shall work from his or her home office, so no travel is required.

Specific requirements for the individual reviewers and the summarizer are listed below[4]. The list is followed by a chronological table, which includes the required steps for the CIE. If any intermediate step in the review process is delayed, all subsequent steps that depend on it will be delayed by an equivalent period.

Individual Reviewers

1. Read and become familiar with the relevant documents provided to the reviewers. The ten age-structured assessments that encompass the focus of the review are presented in 512 pages (of which approximately 1/3 is text).

[4] All reports will undergo an internal CIE review before they are considered final.

2. Discuss the stock assessment with the lead assessment scientist and survey scientists at the AFSC, in Seattle, Washington, from June 11 to June 14, 2007 (see attached agenda).
3. No later than June 29, 2007, submit a written report of findings, analysis, and conclusions. More details on the report outline and organization are provided in Annex I. The report shall be sent via e-mail to Dr. David Die (ddie@rsmas.miami.edu), and to Mr. Manoj Shivlani (mshivlani@rsmas.miami.edu), and to the CIE summarizer (contact information to be provided by the CIE).

Summarizer

7. Become familiar with the documents provided to the individual reviewers.

8. Read the draft reports from the individual CIE reviewers, and draft the summary report following the standard outline. More details on the report outline and organization are provided in Annex I.

9. Complete summary report within one week of receiving final individual reviewer reports from the CIE. The report shall be sent via e-mail to Dr. David Die (ddie@rsmas.miami.edu), and to Mr. Manoj Shivlani (mshivlani@rsmas.miami.edu).

Activity	Deadline
Individual reviewers submit their reports to the CIE and the summarizer.	June 29, 2007
CIE reviews and approves the individual reviewer reports, and provides them to the NMFS COTR.	July 13, 2007
NMFS COTR approves individual reviewer reports.	July 17, 2007
CIE provides final individual reviewer reports to NMFS COTR and to the summarizer.	July 20, 2007
Summarizer provides summary report to CIE.	July 27, 2007

CIE reviews and approves the summary report, and provides it to the NMFS COTR.	August 10, 2007
NMFS COTR approves summary report.	August 14, 2007
CIE provides final summary report to NMFS COTR.	August 17, 2007

Submission and Acceptance of CIE Reports

The CIE shall provide the final reports for review for compliance with this Statement of Work and approval by NOAA Fisheries to the COTR, Dr. Stephen K. Brown (Stephen.K.Brown@noaa.gov), according to the above schedule. The COTR shall notify the CIE via e-mail regarding acceptance of each report. Following the COTR's approval, the CIE shall provide a pdf format version of each approved report to the COTR.

CIE Flatfish Assessment Review
NMFS Alaska Fisheries Science Center
7600 Sand Point Way NE, Building 4
Seattle, Washington
AGENDA May 11 Draft version June 11-14, 2007

Monday June 11th

9:00	Welcome and Introductions, adopt agenda	
9:15	Overview (species, surveys, fishery, catch levels, ABCs, TACs, bycatch)	**Tom**
10:00	Biology (growth, natural mortality, diets, spawning & nursery areas, maturity)	**Buck, Dan, Janet**
11:00	Trawl experiments on herding, escapement, bottom contact issues	**Peter Munro**
11:30	Age Determination of flatfish	**Delsa and Craig**
12:00	Lunch	
13:00	Observer Program	**Jennifer Ferdinand**
14:00	Harvest control rules and projection model	**Jim**
15:00	Summary of on-going research	**Tom, Dan, Janet, Kerim**

Tuesday June 12th

9:00	**Gulf of Alaska and Bering Sea stock assessments**	
	Overview of models	**Jim**
	BSAI Yellowfin sole/northern rock sole	**Tom**
	GOA arrowtooth flounder	**Jack**
	BSAI Greenland turbot	**Jim**
	GOA rex sole	**Jack or Buck**
	GOA shallow-water flatfish	Buck
12:00	*Lunch*	
13:00	**Incorporating ecosystem indicators in stock assessments**	
	Modeling catchability with bottom temperature	**Tom, Paul**
	Catchability and distribution changes with temperature	**Paul**
	OSCURS model to define productivity regimes	**Tom**

	IBM 3-D drift model	**Buck**
Wednesday June 13th		
9:00	Harvest control rules using F_{msy} for yellowfin sole and northern rock sole compared to proxy values	**Tom**
12:00	Lunch	
13:00	Ecosystem aspects of flatfish	**Kerim**
Thursday June 14th		
	Reviewer discussions with assessment authors	

Annex 1: Contents of Reviewer Reports

The following requirements refer to all reports, both the individual reports and the Summary Report.

1. All reports shall be prefaced with an executive summary of findings and/or recommendations.

2. The main body of all reports shall consist of a background, description of review activities, summary of findings, conclusions/recommendations, and references.

3. The reports shall also include as separate appendices the bibliography of all materials provided and any additional papers cited, along with a copy of the statement of work.

Summary report only: The summary report shall include the three individual reviewer reports as appendices.

Appendix 2. Meeting attendees

Alaska Fisheries Science Centre staff

Anne Hollowed
Delsa Anderl
Janet Duffy-Anderson
Craig Castelle
Peter Munro
Jennifer Ferdinand
Paul Spencer
William Stockhausen
Jack Turnock
Jim Ianelli
Tom Wilderbuer
Dan Nichol
Kerim Aydin

CIE Reviewers

Din Chen
Paul Medley
Graham Pilling