Independent Peer Review Report on the

Gulf of Alaska and Bering Sea Flatfish Assessment Review

Prepared for:
The Center for Independent Experts

Dr Kevin Stokes
Stokes.net.nz Ltd
59 Jubilee Rd
Khandallah
Wellington 6035
New Zealand
Ph: +64 (04) 973 7305
E-mail: kevin@stokes.net.nz

EXECUTIVE SUMMARY

General

The Peer Review for Gulf of Alaska and Bering Sea Flatfish took place at the Alaska Fisheries Science Center, Seattle, from 11th to 13th June 2012. The review was hosted by AFSC. Stocks reviewed were from the Gulf of Alaska (GOA) and Eastern Bering Sea (EBS). Four GOA stocks (res sole, Dover sole, northern and southern rock sole) and one EBS stock (yellowfin sole) were reviewed. Terms of Reference (ToR) for the five stocks were variable. There was no chair or rapporteur appointed, with stock assessment authors effectively managing relevant sessions. The meeting was attended solely by AFSC staff as well as the three CIE reviewers. There was no public or other scientific input.

Data availability for the GOA and EBS stocks is markedly different. The EBS yellowfin sole assessment draws on substantial data sufficient to allow state-of-the-art assessment (using an integrated statistical catch age model). Documentation is clear and concise though lacking in some technical (mostly diagnostic) detail. The assessment is robust and represents the best available science. Only one issue was cause for concern. The assessment includes internal fitting of a stock recruit relationship but draws only on data from 1977 or 1978 (unclear) due to a "well known" regime change at that time. The justification for this is not clear in the documentation or the review meeting. More importantly, accepting the date, the relationship is influenced by just a few data points for the years 1977-80. At that time, the stock was recovering from a period of heavy exploitation and was emerging from a different productivity regime. Recruitment was provided by just a few age groups of newly mature fish. It is not clear that the data are representative of stock and recruitment dynamics generally. If the data re excluded, there is no basis for fitting a relationship, and hence for estimating F_{msy} or designating yellowfin sole as a Tier 1 stock. Effectively, the status definition of yellowfin sole is entirely dependent on just a few estimated recruitment values at a time when the stock structure is atypical.

Data availability for the GOA stocks is relatively poor. All of the stocks are very lightly exploited and there are few signals apparent in the raw data which the assessments can use to make inferences about fishing mortality rates. With so little information in the limited data, it is *a priori* unreasonable to expect models reasonably to distinguish fishing mortality variations form demographic noise and observation error. All of the assessments have available an arguably acceptable series of biomass estimates which, together with data-driven estimates of natural mortality, could be used to underpin Tier 5 management. Whether any of the assessments can extract sufficient signals to provide reliable yield per recruit estimates and reliable proxies for B_{msy}/F_{msy} (and hence to underpin Tier 3 management) is debatable.

BACKGROUND

GOA rex sole

Rex sole (*Glyptocephalus zachirus*) is a right-eyed, benthic feeding flatfish distributed from southeastern Alaska to Baja California, at depths from less than 100m to about 800m. The species is widespread and fairly uniformly distributed throughout the GOA, typically in waters between 100m and 200m depth. In the GOA, rex sole live to over 20 years of age, with an observed maximum of 27 years, reaching 50 per cent maturity at 5-6 years and with protracted spawning (October through May). There is little information on rex sole stock structure; the assessment region for rex sole considered in this review covers the whole GOA (although total allowable catch (TAC) and allowable biological catch (ABC) are apportioned at a finer scale based on GOA Groundfish Survey results).

The key terms of reference (ToR) for rex sole, in addition to seeking recommendations on how to improve the assessment, is "…*to evaluate the current approach to determining stock status and future harvest reference points (ABC and OFL)*." The ToR is motivated by differences of opinion between the Plan team and SSC as to appropriate information Tier designation. The Assessment Author and Plan Team's conclusion is that spawner-per-recruit calculations, as required for Tier 3 status definition and setting of F_{OFL} and F_{ABC}, cannot reasonably be made, due primarily to the age at selectivity being so much greater than the age at maturity and the consequent high uncertainty in estimates of any $F_{x\%}$ values. The reported SSC desire is to see the assessment results used and Tier 3 calculations made. The focus of the review for rex sole needs to address this issue.

GOA Dover sole

Dover sole (*Microstomus pacificus*) is a flounder distributed from the Bering Sea and Aleutian Islands to Baja California, with adults found predominantly at depths from 300m to about 1500m. As in other regions, adult (especially female) Dover sole migrate seasonally, spawning from January through August on the shelf and upper slope but moving to shallower waters in summer to feed. In the GOA, Dover sole live to over 50 years of age, with a recently observed maximum of 57 years, reaching 50 per cent maturity at 6-7 years and with protracted (batch) spawning (January through August). No information on Dover sole stock structure or potential southward connection is mentioned in the SAFE Report; the assessment region considered in this review covers the whole GOA (although TAC and ABC are apportioned at a finer scale based on historical catch proportions).

Dover Sole are caught as the predominant species in the deepwater flatfish complex and are taken largely as by-catch in fisheries targeted at halibut, Pacific Cod, and other demersal species. Early catches were exclusively by foreign vessels, joint venture in 1987, and then exclusively domestic vessels from 1988 onwards. Catches in the early (foreign vessel) years declined steadily from 827t in 1978 to just 23t in 1986. While flatfish fisheries generally increased considerably in 1987, the deepwater flatfish catch remained low, with just 56t of Dover sole caught. In 1988, the Dover sole catch increased to over 1,000t and quickly to a high of near 10,000t in 1991 before steadily decreasing to just 400-500t per year in the last decade. Total catches over the past decade have been well below the TAC (of the order of one tenth of the TAC) and discard rates have been high (typically about half). As reported in SAFE 2011 and in the review presentation (Doc 6), annual fishing patterns may vary spatially and temporally (as annual uptake).

The ToR for Dover sole are extensive, covering effectively all assessment data and model components. The motivation for the ToR is that although the assessment in 2011 is just an update of that in 2009, it gives a very different interpretation. The primary difference in model fits can be seen in the estimates of survey selectivity (SAFE, Fig. 5.12 and 5.14-15), consequent interpretations of biomass size, and hence of fishing mortality reference points.

GOA northern and southern rock sole

Rock sole form part of the GOA shallow-water flatfish complex exploited by foreign vessels until 1986, by joint venture vessels in 1987 and by domestic vessels since 1987. Since 1996, rock sole has been split into separate species: northern rock sole (*Lepidopsetta polyxystra*) and southern rock sole (*Pleuronectes bilineata*). The distributions of the two species overlap with northern rock sole found in surveys in the western and central GOA and southern rock sole found substantially in the same area as well as in lower numbers to the eastern GOA. Trends in biomass for the two species are highly correlated (SAFE Fig. 4.1 and 4A.14). Northern rock sole is a late winter spawner, reaching 50% maturity at about 7 years old, while southern rock sole spawns in summer, reaching 50% maturity at 9 years. Maximum age for both species is not noted in the SAFE but appears to be 30 or more from survey age distributions (SAFE Fig. 4.18) and length-at-age plots (SAFE Fig. 4.19). Age at selection (50%) to the fishery (as estimated in the assessment model) is only about 3 or 4 years old for northern rock sole but is about 9 years for southern rock sole, despite similarities in growth pattern for both species until age ten, and generally similar stock and fishery concentrations.

Rock sole are caught as the predominant species in the shallow-water flatfish complex and are both targeted (as flatfish rather than by species) as well as being taken as by-catch in the halibut fishery. By-catch in the halibut fishery has largely constrained the rock sole catches which, as part of the shallow-water flatfish complex, have been well below TAC (SAFE TABLE 4.2). Flatfish catches (including rock sole) in the early (foreign vessel) years declined steadily from over 5,000t in 1978 to below 1,000t in 1986. Catches increased under joint venture fishing in 1987 to 3,561t and have since then fluctuated without clear trend, ranging between about 2,500t and 9,500t with peaks in 1993, 1996, and again in 2008/9. In the last decade, discard rates have been of the order of 6-13%. As reported during the review presentation (Doc 5; also SAFE Figs. 4A.1-2), fishing is concentrated in Area 630 (Central GOA) though surveys indicate the centres of distribution are Area 610 (Western) for northern and southern rock sole, but especially for northern (Doc 5 and SAFE Figs. 4A.5-6). Observer coverage of the fishery and splitting of catches into northern and southern proportions has been low until recent years (SAFE Fig. 4A.3). As fishing patterns may vary spatially and temporally (as annual uptake), and as observer coverage is not designed to be representative and is at a low level (circa 1% per year; SAFE Fig 4A.2), there is considerable uncertainty in catch splitting. Comparing catches to estimates of biomass in surveyed years, extractions have not exceeded 5% in the last two decades and for both northern and southern rock sole have been of the order of 1-3% in the last decade.

The ToR for rock sole are quite general, requiring comment on the analytical approach and data quality, and recommendations for assessment improvements.

BSAI yellowfin sole

Yellowfin sole (*Limandra aspera*) is a small pleuronectid, widely distributed across the northern Pacific from Korea to Canada. On the EBS they are considered as a single stock, with no connection to adjacent areas. Adults migrate between winter spawning grounds near the shelf edge to shallow shelf feeding grounds in summer. They are benthic feeders, mostly on invertebrates. Yellowfin sole live to over 25 years of age, reaching 50 per cent maturity at 10-11 years, by which time they are nearly fully selected to the fishery. L_{inf} for males and females is just 33.7cm and 37.8cm, respectively, with similar length-at-age until about 8 years old. Growth in length is related to spring bottom temperature and growth in weight is additionally correlated to summer bottom temperatures.

Yellowfin sole have been caught in the EBS since the early 1950s and have been targeted initially for foreign vessels, then during the 1980s by joint venture vessels, and, since the late 1980s exclusively by domestic vessels. Catches grew rapidly in the 1950s, exceeding 500,000t in one year in the early 1960s, before declining quickly to average near 120,000 in the 1960s and near 50,000t during the 1970s. With increasing biomass and domestication of the fishery, catches improved, peaking at over 200,000t in the 1980s. In the past decade, catches have risen from a low of 63,000t in 2001 to an average of about 120,000t since 2007.

The fishery is the largest flatfish fishery in the world and has been the subject of intensive monitoring. Catches have been recorded since 1954, with age sampling from 1964; surveys have taken place since 1982, with age and full biological sampling. Environmental data is available from 1982. US observers have been deployed since 1973 on foreign vessels and the domestic fishery has been subject to intense observation and monitoring. Since the passing of Amendment 80 to the BSAI Fisheries Management Plan, monitoring has been extended to include all hauls using modern measuring equipment. Data available on BSAI fisheries are of the highest standard.

The ToR for yellowfin sole require a consideration of the assessment and comment on whether it constitutes best available science. Specific requests also include comment on considerations of fishery rationalisation (i.e., the provision of catch shares) and on the potential use of new survey information from the northern BS, an area now occupied by yellowfin sole with retreating ice cover.

REVIEW PROCESS

General comments

The CIE Flatfish Assessment Review for selected Gulf of Alaska and Bering Sea flatfish stocks took place at the AFSC, Seattle, from 11th to 13th June 2012. Participants in the review are listed in Appendix 3(1). Participants were exclusively AFSC staff, together with three CIE reviewers (Sven Kupschus, Yan Jiao, and Kevin Stokes). There was no chair and, given no requirement for a summary report, no rapporteur was appointed. Materials were provided in advance *via* web links. Prior to the meeting, a Dropbox folder was set up with all files; this was maintained during the meeting.

The Terms of Reference (ToR) for the review are given in Appendix 2, Annex 2. The Terms of Reference (ToR) for each stock are different, each having been prepared by the stock-specific assessment author. For the four GOA stocks, the ToR are aimed at evaluation and provision of advice on various assessment matters. For the single BS stock (yellowfin sole), the ToR are similarly focused but additionally require a clear statement as to whether or not the assessment constitutes the best available science.

The large number of stocks (5), together with variable ToR, appeared initially to be too much reasonably to cover in the short time (3 days) available for the review meeting. The meeting, however, was quite efficient and the presentations reasonably focused. While time was short, it was not unduly constraining for most stocks given the level of detail presented. Nevertheless, there was not time available to allow requests and responses sufficient to delve deeply in to specific issues and with hindsight, more time to explore specific issues would have been beneficial. The meeting followed the outline of the agenda (draft: Appendix 2, Annex 3; final: Appendix 3(2)). The lack of a chair was not generally a problem as each session was effectively managed by the relevant assessment author. However, ensuring consideration of all ToR would have benefited from more formal arrangements.

Reviewer’s role in review activities

The role of the reviewer is set out in the CIE Statement of Work, Attachment A, attached here in Appendix 2, Attachment A. All three CIE reviewers are tasked with producing an independent report to the CIE. The reviewers are not additionally tasked with contributing to any summary or overall report.

Due to illness, and at late notice, I was unable to travel to the meeting. The CIE and AFSC were very helpful in arranging web-based linkage to the meeting. Presentations and teleconference facilities were available throughout and all files were kept up to date and available using Dropbox. The arrangements generally worked well. At times, however, when tightly focused discussions took place around computer screens or printouts it was difficult to participate fully. I am grateful to the CIE and AFSC for making it possible to participate remotely.

SUMMARY OF FINDINGS BY STOCK

GOA rex sole

CIE Reviewers shall evaluate, and make recommendations for improvements on, the current approach to determining stock status and future harvest reference points (ABC and OFL).

As reported in SAFE (2011; Doc 13), Table 6.14, catches are variable, ranging from less than 1,000t per year in the mid-1980s but reaching 4,600t in 1991 and in 2009. Generally, catches are in the range 1,000t to 4,000t. Catches are in general less than half of the TAC and nearer one third of the ABC, consistent with rex sole not being a primary target. Catches are typically of the order of 2 to 4% of the biomass (as estimated from the bottom trawl survey; percentage calculated using Table 6.14). However, while the survey selects fish from about the age at maturity, the fishery selects older fish making comparison difficult. The available survey information suggests biomass has trended upwards since the mid-1980s with relatively little recruitment variability and perhaps an upward trend in recruitment through time (based on examination of Doc 23; also SAFE 2011 Fig. 6.8 and Doc 7). The general picture is of a fishery which only very lightly exploits the resource, with a strong focus on fish much older than the age at maturity.

It appears from the brief description of at-sea catch sampling in the GOA (as described at Doc 4) that it is relatively (to EBS) limited and potentially biased. Fishery age composition data in fact are not used in the assessment, only size composition data with fixed growth parameters. Only triennial and biennial survey age compositions are available. SAFE 2011, Table 6.8, shows the very wide distributions in survey size-at-age. It is noted in the SAFE report that there are area differences in growth rates. There appears to be little information on age past size circa 45cm.

The assessment is effectively forced to fit to the survey biomass trend. Although Markov chain Monte Carlo (McMC) is used, the model does not account for all sources of error, with many key parameters fixed and unexplored and with selectivity assumed constant through time despite changes in design through time (as reflected in an attempt to fit three selectivity blocks in the rock sole assessments). The resulting McMC credible intervals on biomass are very tight. Unsurprising, given the large difference in maturity and fishery selectivity, and the low overall exploitation and lack of contrast in data, fishing mortality and derived estimates of $F_{X\%}$ are poorly defined. SAFE 2011 Fig 6.18 is the best indication of the problem, showing that the quantities of relevance for Tier 3 usage ($F_{35\%}$ and $F_{40\%}$) are not able to be estimated reliably.

The assessment as reported has not been subject to major sensitivity testing, retrospective analyses, likelihood profiling, etc. Nor has there been a systematic, in-depth (and time consuming) consideration of all likelihood components in order to understand tensions between data sets, possible changes in selectivity, etc. A summary of likelihood components is provided at SAFE 2011 Table 6.12 but only of major components. While a more detailed approach to the assessment could be taken, it does not seem warranted given a) the assessment is at least consistent with previous assessments (SAFE 2011 FIG 6.16); b) the fishery is lightly exploited and not targeted; c) it is driven primarily by the bottom trawl survey; d) some age data are due within a year or two. While improved fishery and survey age and growth data (as scheduled) may help improve the assessment, it is also possible that additional data could create tensions between data sets not currently seen.

Given the difficulty of estimating $F_{35\%}$ and $F_{40\%}$, it would be natural to place rex sole in the Tier 5 category on the basis that the assessment cannot provide a reliable estimate of these reference points. TAC could then be set using an estimate or recent average of biomass directly from the survey. Note the assessment, with such high weight placed on the survey, is effectively just a complicated smoother of the survey series; using the assessment-derived B (and $B_{40\%}$) point estimates would be little different to using the survey data directly and transparently. F_{OFL} and $F_{ABC,}$ as per Tier 5 guidance, would be set based on M. Note that a fixed value of M was used in the assessment but there was no consideration in the review of alternative parameterization.

During the review, some time was spent exploring the possibility (suggested by SSC) of using fishery selectivity set equal to maturity as a way of projecting to test for overfished or approaching overfished status, the intention, apparently, to be able to use Tier 3 determination (and the argument being that under the new catch shares regime, targeted fishing could develop). No clear arguments, however, were put forward to support this approach. In principle it is unappealing because as yet there seems to be fundamental lack of information on the fishery processes and appropriate modeling and parameterisation of fishing patterns, or analysis of possible fishing trends in response to catch shares. More could be done to explore historic length-based selectivity.

The Tier system places specific demands on model outputs to support status determination. It is important not to lose sight of the simple conclusions that can be drawn from the catches relative to estimated/observed biomass, for a fishery selecting fish well above the age at 50% maturity, and given available observations. Namely, while the formal Tier 3 status definition cannot be made on technical grounds, it is highly unlikely that rex sole is exploited at a rate in excess of natural mortality or any reasonable definition of F_{OFL}, spawning biomass is at the highest level estimated, 3+ biomass is at or close to the highest level observed/estimated, recruitment has been high in recent times. Overall, although no formal status definition is possible under Tier 3 designation, it is highly unlikely that rex sole is overfished or likely to be in the near future.

GOA Dover sole

CIE Reviewers shall evaluate the current model assumptions and make recommendations for improvements thereof, including:
Use of age data, including:

- *use of age composition data*
- *appropriateness of age range and binning*
- *estimation of size-at-age relationship and variability (external vs. internal to model)*
- *inclusion of ageing error*

Use of size data, including:

- *use of survey size composition data*
- *use of fishery size composition data*

The number and functional forms of estimated selectivity curves, including:

- *fitting different selectivity functions to data from different survey years based on survey depth coverage*
- *types of selectivity curves considered*
- *use of age-based vs. size-based selectivity curves*
- *allowing for annual variability in fishery selectivity*
- *use of size-based selectivity curves for survey data based on trawl net catchability experiments*

Fixing (and updating) the natural mortality rate based on Hoenig, 1983.
Model convergence diagnostics

The assessment, as for rex sole, uses an integrated statistical catch-age model, and attempts to make inferences about fishing mortality rates from various data sources: catches, fishery size compositions, survey biomass estimates and age/size compositions. Growth and natural mortality are fixed and the model can only fit to data by varying annual recruitment, fishery and survey selectivity and fishing mortality rates at age. Selectivity for both the survey and fishery is fit by age for a single time block, although comments in the rock sole SAFE suggest the survey selectivity may better be reflected as three time blocks.

Survey biomass estimates (SAFE, Table 5.5), allowing for assumed availability and with a notional selectivity of 1, range from circa 70,000t to 100,000t. Peak estimates are in the years 1990-93, at the time highest total catches were taken, and again in 2003, following a decade of sustained higher catches. Catches have ranged from a high of about 10% of estimated biomass in the years 1991/92, to less than 1% in the large majority of years, including for the last decade. The generally low level of removals relative to natural mortality suggests *a priori* that the stock has been and is very lightly exploited and unlikely to be overfished. Fitting the data to infer fishing mortality rates is difficult because the low level of exploitation does not create contrast in data that might be detectable against recruitment fluctuations, variations in natural mortality and growth, and sampling error. Given these factors, fishery composition data in particular need to be very well sampled if fishing mortality is to be estimated reliably. As fishery selectivity is fit to age, age sampling and ageing also need to be of high quality and intensity. Neither of these conditions appears to hold. Fishery size composition data for the foreign fishery are not available. Fishery composition data from the 1990s (SAFE Fig. 5.4) appear to be poorly sampled for sex, especially during the high catch period from 1990-1996 (possibly due to difficulty sampling large numbers of young fish in that period?), sampling in 2005-08 was very low and the data are not used in the model. Age data from the fishery appear not to be available and there is no indication that this will change in the near future, with fishery observers not targeting Dover sole which is caught only in small numbers. Overall, while model fitting follows a standard approach, it is hard to see how the age-based model used can be expected to provide credible estimates of fishing mortality. Even if fishery selectivity were fit to size instead of age, the low level of sampling of the very lightly exploited stock suggests that fishing mortality estimation is unlikely to be reliable.

Surveys provide the best source of information on the stock. Surveys have been conducted triennially and then biennially but have in different years extended to different maximum depths. This has been dealt with in a pragmatic and simple manner by using multipliers on availability (SAFE Fig. 5.5). Survey sampling for size and age appears to be good, although 1990 has been dropped from analysis due to skewed sampling (to males). (Note that the data presentation in SAFE Fig. 5.9 should be amended to remove the binned 35-40cm class.) While the data appear to be sound, the primary model fitting difficulty seems to lie in being able to fit age-based selectivity – the 2011 and 2009 models (the same except for data update) show dramatically different parameter fits to the same selectivity function (SAFE Fig. 5.12). The difference is clearly shown in SAFE Figs. 5.12-15 and hence different interpretations of biomass, recruitment and fishing mortality (SAFE Figs. 5.16-23). The 2009 estimated biomass trajectory does not even fall in the confidence intervals on biomass estimated in the 2011 (and barely so for estimated recruitment through time). The assessment tracks the survey observations, effectively smoothing the series. The 2011 assessment closely follows the entire survey series. The 2009 assessment, in contrast, ignores the first two survey points, creating a very different qualitative picture of biomass trends and of current status. It is unclear why the 2009 model does not fit those early observations and the issue was not explored during the review. There is no indication in the SAFE report that early survey (or other) data were treated differently in 2011 and 2009.

During the review the assessment author described the model fitting process. There were no obvious problems with this but the use of a dedicated model application (in AD Model Builder) did not allow quick and detailed exploration of fits and variants (as might be possible e.g. using Stock Synthesis 3 (SS3)). It was not therefore possible in the review meeting to explore in detail why the difference in interpretation occurs between 2009 and 2011 model fits. Only a detailed examination of likelihood components through time might reveal fully how the data influences the fit. It is possible that alternative selectivity functional forms, possibly on size or even blocked in time, might be necessary to stabilise fits and provide credible (explicable) selectivity functions; any such detailed modelling would need to be guided by close examination of likelihoods and by credible arguments to explain how such selectivity changes may have occurred.

[NOTE: Presumably the results in SAFE 2011 reflect model runs for both 2009 and 2011, made in 2011 and with exactly the same assumptions and settings (this was not explicitly noted in the SAFE or review meeting). The difference in fits to the early survey estimates are then very difficult to understand without consideration of fine details of fitting. If, however, the comparison plots are of assessments run in 2009 and then separately in 2011, it would be worthwhile carefully scrutinising input files to check that the 2009 assessment was configured as intended.]

For Dover sole, the fundamental building blocks to enable reliable use of integrated statistical catch-age models seem to be lacking. Namely, reliable fishery size and age sampling and exploitation rates high enough to create signals in the data that can be interpreted against natural variability and given sampling errors. Fishery-independent data appear sound but interpretation is difficult and needs careful consideration. The ToR suggest the intention to adjust the existing model to develop a stable and usable assessment to support Tier 3 designation of Dover sole. A more appropriate approach would be to work from the data to determine what sort of model is possible to allow inferences on biomass and possibly on fishing mortality. Rather than trying to adjust the integrated analyses it might be more profitable to carry out separate analyses of surveys to provide a reliable biomass estimate (which could be used in Tier 5 designation), and of fishery-dependent data to understand gaps, biases and potential usage.

GOA northern and southern rock sole

Evaluation, findings and recommendations of the analytical approach (application of a statistical ADMB integrated catch-age model) used to assess stock status and estimation/presentation of uncertainty.
Evaluation findings and recommendations on quality of input data and methods used to process them for inclusion in the assessment (specifically fishery and survey data).
Recommendations for further assessment improvements for management in both the long and short term.

The assessment uses an integrated statistical catch-age model, and attempts to make inferences about fishing mortality rates from various data sources: annual catches, catch split observations, fishery size compositions, survey biomass estimates and age/size compositions (split by north and south from 1996 onwards). Growth, maturity and natural mortality are fixed by species and sex, and the model can only fit to data by varying annual recruitment, fishery and survey selectivity (by block), fishing mortality rates at age (all by species and sex), and potentially by allowing a differential natural mortality on males. Selectivity for both the survey and fishery is fit variously in model runs by size and age for up to three time blocks, based on brief arguments in SAFE 4.Appendix, Model Specification.

Responding clearly to the ToR is difficult because the SAFE report (including assessment description at 4. Appendix) is hard to follow even with careful reading of all tables and the model specification; the text itself does not clearly explain the model structure or why it has been used, model selection, or the fitting process. Not all Tables or Figures are referred to and there is no detailed analysis of results as part of a model building/selection process. The review meeting did not help fully to understand the modelling or data issues.

The model includes both northern and southern rock sole, by sex, in a mixed fishery. All demographic parameters are by species and sex and selectivities also are by species and sex for the survey and fishery (there is, however, no separation of selectivity by target *versus* bycatch fishery and it is unclear how the separate northern and southern selectivities can be fit to the “early” and “Mid” blocks which are prior to 1996 when species separation was first used (and presumably separate composition data were first available). The only way the species assessments are linked is *via* the total catch and the estimation of the proportion by species and inclusion of the observer-estimated fraction as unspecified, northern or southern rock sole. This is not clear in the text and can be found only in the Appendix, Model Specification, at A.10-15. The data for linking the species within one model are limited (see Table 4A.2 for percentage of catch observed- typically 1% or less by weight and with high inter-annual variability).

It would make good modelling sense to consider simpler formulations of the model(s) before moving to the more complicated one used. Consideration should be given sequentially to a) a combined species model; b) two species models; then c) a linked species model. A combined species model might be useful if there is unreliable sampling of catches and if the species are sufficiently alike to obviate separation. Maturity schedules (as at Table 4.14) and growth (at least for females older than 10 years) appear distinct (as seen in Figs. 4A.21-22; note it would be helpful to see data presented first, not after the results). Whether the differences are sufficient to warrant separation is moot. It would be worthwhile exploring a single model and comparing with separate models first, and then possibly a linked model; for management support purposes a combined model, though biologically wrong, may be simple, reliable and sufficient. Separate models could be run using catches separated by the observer estimates but perhaps with increased observation error on catch by year. It is hard to see why separated models would be less reliable for informing management than the linked model, and it would be easier to consider the assessments separately in detail. The potential advantage of moving to a linked model is that it might be possible to quantify the confidence in the catch separation and to account for it directly in uncertainty measures on quantities of management interest. However, the model does not estimate species fraction and there is therefore no direct comparison available with the observer data to gauge the fit. There is no way of knowing if the linked model is useful or not. Also, given that the model generally under-represents uncertainty, is there any real advantage to be gained by linkage rather than exploring other drivers of uncertainty?

Model selection considered only variations on selectivity blocking and natural mortality offset estimation. The description at SAFE pp. 452-453 is difficult to follow without reference to details contained in multiple tables. The explanation for including the offset in natural mortality is to allow fitting of an observed high female fraction in surveys (SAFE Fig. 4A.15). Use of Model 1, including the male offset and estimating M at 0.26 compared to female M of 0.2, does result in lower LL than Model 2 (with no offset), with all gain in the survey fraction female likelihood component. However, it is clear from the figure that even with the male offset in M, the model cannot capture the observed survey fraction. It seems likely therefore that whether or not there is a difference in M by sex, the skewed sex ratio is more likely a function of survey timing and location or sex-specific selectivity in the survey. Modelling with a large offset in male M may not be appropriate if it "corrects" for the survey sex ratio but that ratio is itself a misrepresentation of the population sex ratio. Is it possible to compare sex ratios for recent surveys with fishery data, both in the shallow-water flatfish fishery but also in by-catch fisheries?

Selectivity blocking is very briefly explained late in the document (p. 542). Rather than pre-defining blocks, it would be instructive to fit to a single block and examine the likelihoods and other diagnostics by year to look for break points in fits. If those breakpoints were consistent with rational explanations there would be greater support for maintaining them. Currently the size-based selectivities by block are unconvincing. For northern rock sole, the fishery and all survey period selectivities are similar. For southern rock sole, however, it is unclear why there is such a big difference between fishery and survey selectivities (though spatial and temporal coverage with respect to fish distribution may be a factor). What is clear is that the variability between blocks is high and unlikely to be credible. Examining the likelihoods in table 4A.5 for Models 1, 3 and 6, not fitting size-based selectivity to the early period makes little or no difference to fits to survey length compositions. The main reason for the small increase in overall likelihood is the increase in the fit to fraction female in the survey. When the middle selectivity block is also not fit, the effect on the fit to female fraction is lost, the major effect is on the fit to southern rock sole survey age composition data and also to unspecified length and length-at-age fits. Overall, there appears to be little information on the early period southern rock sole selectivity while the information on the middle period selectivity is real but will require careful examination and referencing to survey changes before it is credible.

As for rex and Dover sole, rock sole are lightly exploited. The catches relative to raw survey-derived biomass estimates suggest a very low fishing mortality rate. The assessment, consistent with the catch and biomass estimates, suggests a fishing mortality rate about one quarter or less of natural mortality (SAFE Figs. 4A12-13). Fishery sampling is poor and there should be little expectation of information on fishing mortality by age or recruitment in the data, especially given the confounding factor of species splitting (again poorly sampled). The survey age composition data (presented in Doc 5) do not apparently show clear cohort structure, though there is perhaps some indication of a signal for northern rock sole for a 1999/2000 cohort which may be reflected in the relatively strongly estimated 1999 YC in Fig. 4A.9. For southern rock sole, Fig. 4A.9 shows a number of strongly estimated YC; those prior to 1990 are possibly indicated in the survey age composition data but the strongly estimated 1998 and 2003 YCs are not at all apparent in the raw data. Generally, the survey information does not appear to be highly informative and the unexplained high fraction of females and poor selectivity fits (especially for southern rock sole) cause some concern.

BSAI yellowfin sole

Evaluation of the analytical approach (application of a statistical ADMB integrated catch-age model) and model assumptions used to assess stock status and stock productivity.
Evaluation of the implications of using the Northern Bering Sea research results as an index of abundance if yellowfin sole increasingly occupy this area with changing climate.
Determination of whether the assessment represents the best available science for the stock assessment of BSAI yellowfin sole, including considerations of fishery rationalization on timing and selectivity of fishery.

The assessment uses an integrated statistical catch-age model, and attempts to make inferences about fishing mortality rates from various excellent data sources: annual catches, fishery age compositions, survey biomass estimates and age compositions, as well as using direct estimates of weight-at-age by year from the survey and environmental data. The model estimates stock recruitment parameters internally. Alternative growth models and natural mortality estimation were explored but the direct weight-at-age estimates and fixed M were used in the final assessment. Alternative weight-at-age models have been considered but use of direct weight estimates is appropriate given the data quality/availability, and the annual variability in growth due to complex temperature correlations. Use of fixed M seems well justified given past analyses and profiling, and the potential to confound M estimation in the annually estimated survey catchability parameter. The assessment document is clear (if lacking in diagnostics, sensitivities, etc.), and the assessment appears to be well fit to data and credible.

The assessment is unusual in two respects. First, because of the extensive and excellent data available, selectivity is fit by sex and age for every year. Second, given the strong relationship between temperature and catchability, this is estimated annually for the survey. A third issue of note is the estimation of stock recruit parameters (for the Ricker relationship) in which a curtailed data set is used, affecting F_{msy} and B_{msy} estimation. The estimation of the stock recruit relationship allows designation of yellowfin sole as Tier 1.

Stock-recruit fits are shown for the whole data period and for data to and from 1978 (according to SAFE Fig. 4.12 upper panel). The text at p. 591 suggests that the periods are 1955-1977 and from 1978. The point labels in the lower panel of Fig. 4.12 suggest the final fit is for the period 1977 onwards. Text on p. 594 suggests the period used is from 1978 onwards. There is a lack of consistency in the description/labels. Use of the fit as shown in the lower panel of the figure does seem most consistent with the explanation given for its use (p. 592) – that there was a "well documented regime shift in 1977". The documentation is clearly a continuation of past documents and is not comprehensive. For example, it lacks diagnostics and model development details. Overall, however, it is clear and excellent. Given the importance to management if deciding on a period for stock-recruit fitting, it would be helpful to have more detail and explanation to justify the choice of years from 1977 onwards. During the review, the issue was raised and the answer was effectively a repeat of the statement in the SAFE – it is "well known". This is not convincing. Especially concerning is that and stock-recruit signal will likely be confounded in the early years of the new "regime" as it was a time when the biomass was increasing and the age structure filling in rapidly following a period of heavy exploitation (under a presumed different productivity regime). The age composition plots (Doc. 31, slide 6) show clearly the limited number of mature females in 1977-80 and the slow filling out of the age structure over the next decade. The stock structure in 1977-80 is a carry-over from the heavy exploitation during a higher productivity regime and the S-R "data" may not be good indicators of the stock recruit relationship pertinent to the defined regime period. It is problematic that those few data points entirely influence the stock-recruit fit. Overall, the reliability of the stock recruit fit, and hence of F_{msy}, B_{msy} estimation, must be in doubt.

[NB There seems to be an error at Table 4.13 and/or Fig. 4.13. It appears that in Fig. 4.13, Models A and B have been reversed. This would be consistent with the numbers in Table 4.13 as well as with Fig 4.20.]

With respect to fitting annual survey catchability (using two parameters by relating it to temperature), the estimation of q effectively does two separate things. First, it provides an estimate of average biomass scaling (through the parameter alpha). The scaling makes sense (as argued in terms of herding) but it is unclear where the model draws information sufficient to permit the estimation. Keeping M constant helps in determining alpha but if M were to be estimated there would be a confounding with the alpha parameter. Second, the beta parameter in the q estimation relates catchability to temperature. It is this that improves the detailed fit to the biomass indices, enabling better tracking of the biomass signal (as at SAFE Fig.4.18) but it is unclear what benefit this creates to determination of status or in forecasting.

The annual catch patterns for 2007 onwards are shown at Fig. 4. 1. Years since 2008 are fairly consistent, all showing a fast initial uptake and a clear plateau mid-year (associated with poor quality fish). All four years are different to the 2007 pattern. This is at least suggestive that rationalisation has led to more measured and evenly spread fishing. Whether it has resulted in changed selectivity can be considered with reference to the estimated selectivities using the assessment. Selectivity is fit by year for each sex. This is only possible with such good age data from the fishery garnered by extensive and intensive sampling. The results are impressive and provide a useful picture of the fishery through time (Doc.31, slides 25-26). It is not obvious, however, at least as yet, if the rationalisation has led to any fundamental change in selectivity, or whether any account needs to be taken of this in projections. In fact, it is unclear on p. 595 exactly how selectivity is used to make projections. For the projection to the next year presumably the most recent selectivity functions for males and females are used. This is not explicit. Neither is it made clear what selectivity assumptions are used for subsequent years.

In any case, regarding the ToR, it does not appear to be necessary to make any changes to selectivity uses in projections. As data are accumulated and the assessment updated, selectivity estimates will accrue and will feed in to projections. If selectivity stabilises with a new pattern, this will incrementally be adapted to. Only if rationalisation were to lead to such changes in selectivity that reference points were affected, might it be necessary to reconsider the current approach.

The ToR require an evaluation of the implications of using the Northern Bering Sea research results as an index of abundance if yellowfin sole increasingly occupy this area with changing climate. During the review just one slide was provided that showed the Northern BS area in which one survey has taken place and in which yellowfin sole (substantially juvenile) were observed in abundance. Currently, with no details of linkage between the fish in the NBS and EBS, and only one survey, it is not sensible to try to incorporate the information in to an abundance index or to modify the assessment model. If, as expected, the northern BS becomes available to species as ice cover reduces, then it may be that cold tolerant species such as yellowfin sole expand their distribution. So long as that expansion is additional to the current distribution, and so long as no fishing takes place, then ignoring data from the region can only be precautionary in terms of management. If, however, expansion accompanies reduction from the historical area of distribution, if migration patterns change, if recruitment processes were affected, if growth and mortality were affected (directly or indirectly), or if fishing were to expand, etc…, then continuous monitoring of the NSB is important. Only through regular surveys might it be possible to understand the linkage between species in the EBS and NBS and to see any changes in biological characteristics. At this stage, it would be premature to contemplate specific actions other than monitoring with a view to reviewing information in a few years and then to consider whether stock definitions, assessment (and management) would require modification. During the review, "experimental fishing" was raised as a possibility. If the stock (and other stocks) do expand to the north, fishing may become a desirable activity. However, until more is understood of the biological and ecological processes involved, this could be premature. It is not just a matter of species X expanding in to area Y. The whole ecology of the region will be changing and there is no clear concept of what will constitute the new normal, including appropriate fishery options.

CONCLUSIONS AND RECOMMENDATIONS

GOA rex sole

The stock assessment of rex sole uses standard methods and makes a good attempt at inferring management quantities from limited data containing little information. Improved and extended fishery and survey age data are expected in the next year and it will of course be interesting to attempt further analyses, possibly taking account of spatial differences in growth and fishing pattern, and perhaps considering length-based selectivity processes. The problems inherent in the assessment, however, may not be simply resolved. More data does not necessarily mean more information from which to extract signals. Further, more data can create conflict between datasets.

The Tier system places specific demands on model outputs to support status determination. It is important not to lose sight of the simple conclusions that can be drawn from the catches relative to estimated/observed biomass, for a fishery selecting fish well above the age at 50% maturity, and given observations. Namely, while the formal status definition cannot be made, it is highly unlikely that rex sole is exploited at a rate in excess of natural mortality, spawning biomass is at the highest level estimated, 3+ biomass is at or close to the highest level observed/estimated, recruitment has been high in recent times. Overall, although no formal status definition is possible under Tier 3 designation, it is highly unlikely that rex sole is overfished.

With regard to Tier designation, information exists to designate rex sole as Tier 5. Direct use of survey estimates of biomass, or perhaps estimates derived from the assessment, could be used, together with the simple characterization of M. Logically, it would be sensible to use the survey-derived estimate of biomass unless and until there is more confidence in the assessment. More work could be done to refine the estimate of M but it is unclear if this is important or not. It is difficult to argue for Tier 3 designation. The assessment fails to provide a reliable point estimate of F35% and F40%. There is no rational basis for inventing selectivity and exploitation rate scenarios to test whether rex sole is overfished or approaching an overfished state.

GOA Dover sole

The stock assessment of Dover sole uses standard methods and makes a good attempt at inferring management quantities from limited data containing little information. Apart from standard fishery and survey data updates, no new data are expected soon. Attempts to understand why the 2009 and 2011 assessments differ will therefore need to concentrate on examination of modeling details using a range of approaches (examination of detailed likelihood components, profiling, etc.). The problems inherent in the assessment, however, may not be simply resolved. The tools to hand (integrated statistical catch-age models) are powerful but if the data do not contain sufficient information, reliable estimation of management quantities (for Tier 3) may simply not be feasible.

The Tier system places specific demands on model outputs to support status determination. It is important not to lose sight of the simple conclusions that can be drawn from the catches relative to estimated/observed biomass, for a fishery selecting fish well above the age at 50% maturity and with protracted life span, and given observations. Namely, while the formal status definition cannot be made, it is highly unlikely that Dover sole is exploited at a rate in excess of natural mortality (or any rational F_{OFL}), and spawning biomass is highly likely to be well above B_{msy} or proxies such as $B_{35\%}$. Overall, although no formal status definition is possible under Tier 3 designation, it is highly unlikely that Dover sole is overfished.

With regard to Tier designation, as for rex sole, information exists to designate Dover sole as Tier 5. Direct use of survey estimates of biomass could be used, together with the simple characterization of M. It is difficult to argue for Tier 3 designation.

GOA northern and southern rock sole

The stock assessment of northern and southern rock sole uses standard methods in an attempt at inferring management quantities from limited data containing little information for two linked stocks where the linking fishery (not biological) information is poor. Apart from standard fishery and survey data updates, no new data are expected soon. Attempts to explore and understand the data and possible assessments are therefore limited to what exists now. The current assessment represents a considerable amount of effort but the report does not fully explain what has been done and why, and how it has been interpreted. This may well be due to time constraints having undertaken so much work. The report currently is poorly structured and is not convincing. Rather than working from the current multi-stock assessment and trying to refine it, it would likely be more profitable to explore externally the fishery and survey data sources and to build systematically from a simpler starting point (combined species model, then separate species models, then perhaps linked model), examining in detail model diagnostics and only adding complexity (in selectivity fitting, natural mortality estimation, model structure…) as necessary and as justified.

As noted above for other stocks, the tools to hand (integrated statistical catch-age models) are powerful but if the data do not contain sufficient information, reliable estimation of management quantities (e.g. for Tier 3) may not be feasible.

BSAI yellowfin sole

The stock assessment of yellowfin sole uses standard methods and, given exceptional data availability, provides an unusually good base for management decision making. The science is the best available. The documentation is clear and concise but as a continuation of many years of effort and reporting is not complete in some places. In particular, for a stock assessment, many technical details and diagnostics are missing. This is not a major issue given the quality of the work and ability to refer to previous materials. The assessment is essentially straightforward though it includes some unusual features. These include the inclusion of temperature to allow survey catchability estimation by year. This is simple enough, cleans up the fits slightly and requires few parameters. It is not obvious, however, that it provides any advantage and could create problems if natural mortality estimation were considered in the future. The most notable feature of the assessment, as for some other EBS assessments and due to the unusual data availability, is the estimation of fishery selectivity by year and species. This allows a detailed look at fishery functioning. Although it is too soon yet to judge how fishery rationalisation has affected fishing patterns, and whether this needs to be accounted for in the assessment and projections, the ability to estimate selectivity annually will provide a powerful tool for understanding change and modifying projections if necessary.

One area of the assessment is of concern. The assessment estimates the stock recruit relationship internally and uses it to estimate a *pdf* of F_{msy} and other quantities for use in Tier 1 management. On the surface, this is good. However, the way data are split on the basis of a known (but unexplained/undocumented) regime shift is a concern. There are some minor discrepancies in the dates described at various places but the major concern must be the influence due to just a few estimated recruitment values in 1977-80. They drive the analysis and the F_{msy} estimation (and hence any advice). Those recruitment estimates, however, are associated with low biomasses at the start of the defined regime period when the stock was recovering from a low level and restricted age range, with few fish over the age at 50% maturity. Although the fitting is robust and the resultant F_{msy} reliable technically, it is not clear that the stock recruit relationship and hence F_{msy} are truly reliable given a) the inclusion of atypical biomass points, and b) the high influence of each of those points. It is recommended to reconsider the stock recruit fitting with or without the 1977 point (it is unclear if it should be included), testing the influence of each of the 1977, 1978, 1979 points, and with regard to the validity/meaning of inclusion of biomass at the start of the period. It is noted that if those points are excluded, there would be no basis in data to fit a stock recruit curve – implying that Tier 1 status might not be appropriate.

With regard to NBS expansion, it is recommended at this time to keep monitoring with a view to considering implications when more data have been collected. It is recommended at this stage not to modify the assessment or to consider exploratory fishing.

APPENDIX 1

BIBLIOGRAPHY

Prior to the Workshop, materials relevant to all terms of reference were made available *via* web links. During the workshop multiple presentations were given, and additional materials were provided on request, including further background documents and presentations as well as responses to Panel requests. All files were made available at the AFSC on a server which was accessed using using AFSC guest guest Wi-Fi throughout the meeting. All files (see table below) were additionally made available on a Dropbox folder managed by Jim Ianelli of the AFSC.

Files available during meeting and on Dropbox folder		
Doc	**File name**	**Description**
	Agenda for 2012 flatfish CIE (1).docx	Final meeting agenda
1	BSAIintro.pdf	BSA SAFE (2011) introductory section
2	BSAIyfin.pdf	BSA Yellowfin stock Assessment document (from BSA SAFE, 2011)
	CIE agenda for flatfish.doc	Draft meeting agenda (doc)
	CIE agenda for flatfish.pdf	Draft meeting agenda (pdf)
	CIE Flatfish assessment SOW.doc	CIE Flatfish assessment SOW (doc)
3	CIE Flatfish assessment SOW.pdf	CIE Flatfish assessment SOW (pdf)
4	CIE Review NPGOP_slides.pdf	Presentation (Lisa Thompson) – North Pacific Groundfish Observer Program (pdf)
	CIE Review NPGOP_slides.pptx	Presentation (Lisa Thompson) – North Pacific Groundfish Observer Program (pptx)
5	CIEreview.GOArs.2012.pdf	Presentation (Teresa A'mar) – CIE Review of the GOA northern and southern rock sole stock assessment (pdf)
	CIEreview.GOArs.2012.ppt	Presentation (Teresa A'mar) – CIE Review of the GOA northern and southern rock sole stock assessment (ppt)
	ciereviewoftheebsbottomtrawlsurvey.zip	Folder containing (3) reports by CIE reviewers on the EBS crab and groundfish bottom trawl surveys
6	CIEReview_DoverSole.pdf	Presentation (William Stockhausen) – Dover Sole (pdf)
	CIEReview_DoverSole.pptx	Presentation (William Stockhausen) – Dover Sole (pdf)
7	CIEReview_GOARexSole.pdf	Presentation (William Stockhausen) – GOA Rex Sole (pdf)
	CIEReview_GOARexSole.pptx	Presentation (William Stockhausen) – GOA Rex Sole (pdf)
8	DoverSole_AgeCompsData.xlsx	Dover Sole age composition data provided during meeting
9	DoverSole_DiffAgeCompComparisonsFemale.pdf	Dover Sole residual plots by cohort (female) provided during meeting
10	DoverSole_DiffAgeCompComparisonsMale.pdf	Dover Sole residual plots by cohort (male) provided during meeting
	FlatFishCIE_2012.htm	Html copy of draft agenda
11	Flatfish_A&G_CIE review.pdf	Presentation (Tom Helser) – Flatfish age determination at AFSC(pdf)
	Flatfish_A&G_CIE review.pptx	Presentation (Tom Helser) – Flatfish age determination at AFSC(pptx)
12	GOA2012SAFE_DeepwaterFlatfish.pdf	Deepwater Flatfish stock assessment document (from GOA SAFE, 2011)
13	GOA2012SAFE_RexSole.pdf	Rex Sole stock assessment document (from GOA SAFE, 2011)
14	GOAAOFlatfishCIEpalsson.pdf	Presentation (RACE Division) – GOA and AI Bottom Trawl Survey (pdf)
	GOAAOFlatfishCIEpalsson.pptx	Presentation (RACE Division) – GOA and AI Bottom Trawl Survey (pptx)
15	Haynie Abbott CIE A80.pdf	Presentation (Alan Haynie) – Rationalization of the BSAI Amendment 80 Fleet (pdf)
	Haynie Abbott CIE A80.pptx	Presentation (Alan Haynie) – Rationalization of the BSAI Amendment 80 Fleet (pptx)
16	Nichol_CIE_6_11_2012.pdf	Presentation (RACE Division) – EBS shelf Bottom Trawl Survey of groundfish and invertebrate resources(pdf)
	Nichol_CIE_6_11_2012.pptx	Presentation (RACE Division) – EBS shelf Bottom Trawl Survey of groundfish and invertebrate resources(pptx)
	Notes for Dover sole and assmnts.docx	J. Ianelli partial notes for Dover Sole and northern and southern rock (made during meeting)
17	Nov_2011_BSAI_Minutes.pdf	Minutes from BSAI Plan team meeting, November 2011
18	Nov_2011_GOA_Minutes.pdf	Minutes from GOA Plan team meeting, November 2011
19	Nov_2011_JPT_Minutes.pdf	Minutes from joint BSAI and GOA Plan team meeting, November 2011
20	overview.ppt	Presentation (Tom Wilderbuer) – Overview of Flatfish Fisheries and Management (ppt)
21	rex sole survey-at-age.xlsx	Rex Sole survey residual plots by cohort provided during meeting
22	RexSoleAgeCompsData.xlsx	Rex Sole age composition data provided during meeting

23	RexSoleSurvey.docx	Diagnostic plots provided by Sven Kupschus during meeting
24	RexSole_DiffAgeCompComparisonsFemale.pdf	Rex Sole residual plots by cohort (female) provided during meeting
25	RexSole_DiffAgeCompComparisonsMale.pdf	Rex Sole residual plots by cohort (male) provided during meeting
26	SomertonEtAl2007_FishBull.pdf	Somerton, Munro and Weinberg (2007) Whole-gear efficiency of a benthic survey trawl for flatfish. *Fish. Bull. 105:278-291*
27	SSC Dec2011Draft.pdf	Draft SSC Report to the NPFMC, December 2011
28	SUMMARY_2011_BSAI.xls	Summary information on BSAI fleet/gear structure provided during meeting
29	SUMMARY_2011_GOA.xls	Summary information on GOA fleet/gear structure provided during meeting
30	yfs age comps.xlsx	Yellowfin Sole age composition data provided during meeting
31	yfs presentation to cie panel.pptx	Presentation (Tom Wilderbuer) – CIE Review of the Bering Sea/Aleutian Island yellowfin sole stock assessment
	Other references	

APPENDIX 2

Attachment A: Statement of Work for Dr. Kevin Stokes

External Independent Peer Review by the Center for Independent Experts

Peer Review of the BSAI and GOA flatfish stock assessments

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

Project Description: The Alaska Fisheries Science Center (AFSC) requests a Center of Independent Experts (CIE) review of 4 Bering Sea/Aleutian Islands (BSAI) and Gulf of Alaska (GOA) flatfish stock assessments. They include: GOA northern and southern rock sole, GOA Dover sole, GOA rex sole and BSAI yellowfin sole. The BSAI and GOA flatfish resources are large, subject to significant fisheries and are key components of the BSAI and GOA ecosystems. The flatfish stock assessments routinely undergo thorough review by the AFSC, the North Pacific Fisheries Management Council's Groundfish Plan Teams and Scientific and Statistical Committee, and members of the public. However, the BSAI and GOA flatfish stock assessments have not had the benefit of a CIE review since 2007. Since 2007, several modifications to existing assessment and projection models have been implemented, and a new assessment for Gulf of Alaska northern and southern rock sole has been developed. These innovations have not been reviewed by the CIE. The Alaska Fisheries Science Center desires an independent peer review of these stocks to assess the quality of the assessments and to ensure that the North Pacific Fishery Management Council bases its decisions on the best available information. Therefore, a CIE review in 2012 would be timely. The Terms of Reference (ToRs) of the peer review are attached in **Annex 2**. The tentative agenda of the panel review meeting is attached in **Annex 3**.

Requirements for CIE Reviewers: Three CIE reviewers shall conduct an impartial and independent peer review in accordance with the SoW and ToRs herein. CIE reviewers shall have working knowledge and recent experience in the application of stock assessment, including population dynamics, survey methodology, estimation of parameters in complex nonlinear models, and the AD Model assessment program in particular. Reviewers should also have experience conducting stock assessments for fisheries management. Each CIE reviewer's duties shall not exceed a maximum of 14 days to complete all work tasks of the peer review described herein.

Location of Peer Review: Each CIE reviewer shall conduct an independent peer review during the panel review meeting scheduled in Seattle, Washington tentatively during June 11-13, 2012.

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

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

Foreign National Security Clearance: When CIE reviewers participate during a panel review meeting at a government facility, the NMFS Project Contact is responsible for obtaining the Foreign National Security Clearance approval for CIE reviewers who are non-US citizens. For this reason, the CIE reviewers shall provide requested information (e.g., first and last name, contact information, gender, birth date, passport number, country of passport, travel dates, country of citizenship, country of current residence, and home country) to the NMFS Project Contact for the purpose of their security clearance, and this information shall be submitted at least 30 days before the peer review in accordance with the NOAA Deemed Export Technology Control Program NAO 207-12 regulations available at the Deemed Exports NAO website: http://deemedexports.noaa.gov/sponsor.html).

Pre-review Background Documents: Two weeks before the peer review, the NMFS Project Contact will send (by electronic mail or make available at an FTP site) to the CIE reviewers the necessary background information and reports for the peer review. In the case where the documents need to be mailed, the NMFS Project Contact will consult with the CIE Lead Coordinator on where to send documents. CIE reviewers are responsible only for the pre-review documents that are delivered to the reviewer in accordance to the SoW scheduled deadlines specified herein. The CIE reviewers shall read all documents in preparation for the peer review.

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

Contract Deliverables - Independent CIE Peer Review Reports: Each CIE reviewer shall complete an independent peer review report in accordance with the SoW. Each CIE reviewer shall complete the independent peer review according to required format and content as described in Annex 1. Each CIE reviewer shall complete the independent peer review addressing each ToR as described in Annex 2.

Other Tasks – Contribution to Summary Report: Each CIE reviewer may assist the Chair of the panel review meeting with contributions to the Summary Report, based on the terms of reference of the review. Each CIE reviewer is not required to reach a consensus, and should provide a brief summary of the reviewer's views on the summary of findings and conclusions reached by the review panel in accordance with the ToRs.

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

1) Conduct necessary pre-review preparations, including the review of background material and reports provided by the NMFS Project Contact in advance of the peer review.
2) Participate in the panel review meeting in Seattle, Washington during June 11-13, 2012.
3) In Seattle, Washington during June 11-13, 2012 as specified herein, conduct an independent peer review in accordance with the ToRs (**Annex 2**).
4) No later than July 9, 2012, each CIE reviewer shall submit an independent peer review report addressed to the "Center for Independent Experts," and sent to Manoj Shivlani, CIE Lead Coordinator, via email to shivlanim@bellsouth.net, and CIE Regional Coordinator, via email to David Die ddie@rsmas.miami.edu. Each CIE report shall be written using the format and content requirements specified in Annex 1, and address each ToR in **Annex 2**.

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

May 1, 2012	CIE sends reviewer contact information to the COTR, who then sends this to the NMFS Project Contact
June 1, 2012	NMFS Project Contact sends the CIE Reviewers the pre-review documents
June 11-13, 2012	Each reviewer participates and conducts an independent peer review during the panel review meeting
July 9, 2012	CIE reviewers submit draft CIE independent peer review reports to the CIE Lead Coordinator and CIE Regional Coordinator
July 23, 2012	CIE submits CIE independent peer review reports to the COTR
July 30, 2012	The COTR distributes the final CIE reports to the NMFS Project Contact and regional Center Director

Modifications to the Statement of Work: This 'Time and Materials' task order may require an update or modification due to possible changes to the terms of reference or schedule of milestones resulting from the fishery management decision process of the NOAA Leadership, Fishery Management Council, and Council's SSC advisory committee. A request to modify this SoW must be approved by the Contracting Officer at least 15 working days prior to making any permanent changes. The Contracting Officer will notify the COTR within 10 working days after receipt of all required information of the decision on changes. The COTR can approve changes to the milestone dates, list of pre-review documents, and ToRs within the SoW as long as the role and ability of the CIE reviewers to complete the deliverable in accordance with the SoW is not adversely impacted. The SoW and ToRs shall not be changed once the peer review has begun.

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

Applicable Performance Standards: The contract is successfully completed when the COTR provides final approval of the contract deliverables. The acceptance of the contract deliverables shall be based on three performance standards:
(1) each CIE report shall completed with the format and content in accordance with **Annex 1**,
(2) each CIE report shall address each ToR as specified in **Annex 2**,
(3) the CIE reports shall be delivered in a timely manner as specified in the schedule of milestones and deliverables.

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

Support Personnel:

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

Manoj Shivlani, CIE Lead Coordinator
Northern Taiga Ventures, Inc.
10600 SW 131st Court, Miami, FL 33186
shivlanim@bellsouth.net   Phone: 305-383-4229

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

Key Personnel:

Tom Wilderbuer, NMFS Project Contact
NMFS Alaska Fisheries Science Center
7600 Sand Point Way NE., Seattle, WA 98115-6349
tom.wilderbuer@noaa.gov   Phone: 206-526-4224

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

1. The CIE independent report shall be prefaced with an Executive Summary providing a concise summary of the findings and recommendations, and specify whether the science reviewed is the best scientific information available.

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

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

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

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

   d. Reviewers shall provide a critique of the NMFS review process, including suggestions for improvements of both process and products.

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

3. The reviewer report shall include the following appendices:

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

Annex 2: Tentative Terms of Reference for the Peer Review

2012 CIE Review for selected Gulf of Alaska and Bering Sea flatfish

GOA Dover sole TOR

CIE Reviewers shall evaluate the current model assumptions and make recommendations for improvements thereof, including:
Use of age data, including:

- use of age composition data
- appropriateness of age range and binning
- estimation of size-at-age relationship and variability (external vs. internal to model)
- inclusion of ageing error

Use of size data, including:

- use of survey size composition data
- use of fishery size composition data

The number and functional forms of estimated selectivity curves, including:

- fitting different selectivity functions to data from different survey years based on survey depth coverage
- types of selectivity curves considered
- use of age-based vs. size-based selectivity curves
- allowing for annual variability in fishery selectivity
- use of size-based selectivity curves for survey data based on trawl net catchability experiments

Fixing (and updating) the natural mortality rate based on Hoenig, 1983.
Model convergence diagnostics

GOA rex sole TOR

CIE Reviewers shall evaluate, and make recommendations for improvements on, the current approach to determining stock status and future harvest reference points (ABC and OFL).

GOA northern and southern rock sole TOR

1. Evaluation, findings and recommendations of the analytical approach (application of a statistical ADMB integrated catch-age model) used to assess stock status and estimation/presentation of uncertainty.

2. Evaluation findings and recommendations on quality of input data and methods used to process them for inclusion in the assessment (specifically fishery and survey data).
3. Recommendations for further assessment improvements for management in both the long and short term.

BSAI yellowfin sole TOR
Evaluation of the analytical approach (application of a statistical ADMB integrated catch-age model) and model assumptions used to assess stock status and stock productivity.
Evaluation of the implications of using the Northern Bering Sea research results as an index of abundance if yellowfin sole increasingly occupy this area with changing climate.
Determination of whether the assessment represents the best available science for the stock assessment of BSAI yellowfin sole, including considerations of fishery rationalization on timing and selectivity of fishery.

Annex 3: Tentative Agenda

CIE Flatfish assessment review

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

Article I. AGENDA *JANUARY 20 DRAFT VERSION* June 11-13, 2012

Monday June 11th

9:00	Welcome and Introductions, adopt agenda	**Sandra**
9:15	Overview (species, biology, surveys, fishery, catch levels, ABCs, TACs, bycatch)	**Tom**
10:00	Bering Sea trawl survey	**RACE Division**
10:30	Gulf of Alaska trawl survey	**RACE Division**
11:00	Coffee break	
11:20	Observer Program	**FMA Division**
11:50	Age Determination	**Delsa and Beth**
12:30	Lunch	
1:30	Effect of rationalization on flatfish fisheries	**REFM Economic subtask**
2:30	GOA rex sole	**Buck**

Tuesday June 12th

9:00	Gulf of Alaska Dover sole	**Buck**
11:00	Coffee break	
11:20	Gulf of Alaska Dover sole (continued)	**Buck**
12:30	Lunch	
1:30	Gulf of Alaska northern and southern rock sole	**Teresa**

Wednesday June 13th

9:00	Bering Sea yellowfin sole	**Tom and Jim**
11:00	Coffee break	
11:20	Bering Sea yellowfin sole (continued)	**Tom and Jim**
12:30	Lunch	
1:30	CIE panel discussion (assessment authors will be available)	

APPENDIX 3
PERTINENT INFORMATION FROM THE REVIEW

1) Participants

Sven Kupschus	CEFAS, UK
Yan Jiao	Virginia Tech
Kevin Stokes	stokes.net.nz Ltd, New Zealand
Anne Hollowed	AFSC Status of stocks
Buck Stockhausen	AFSC Status of stocks
Teresa A'mar	AFSC Status of stocks
Tom Wilderbuer	AFSC Status of stocks
Sandra Lowe	AFSC Status of stocks
Jim Ianelli	AFSC Status of stocks
Loh-Lee Low	AFSC International coordination
Alan Haynie	AFSC Economics program
Dan Nichol	AFSC Bering Sea survey program
Wayne Palsson	AFSC Gulf of Alaska survey program
Tom Helser	AFSC Age and growth program
Lisa Thompson	AFSC Observer program

2) Final Agenda

CIE Flatfish assessment review

NMFS Alaska Fisheries Science Center
Marine Mammal Conference Room
7600 Sand Point Way NE, Building 4
Seattle, Washington

Article II. AGENDA *FINAL VERSION* June 11-13, 2012

Monday June 11th

9:00	Welcome and Introductions, adopt agenda	**Sandra**
9:15	Overview (species, biology, surveys, fishery, catch levels, ABCs, TACs, bycatch)	**Tom**
10:00	Bering Sea trawl survey	**Dan Nichol**
10:30	Gulf of Alaska trawl survey	**Wayne Palsson**
11:00	Coffee break	
11:20	Observer Program	**Lisa Thompson**
11:50	Age Determination	**Tom Helser**
12:30	Lunch	
1:30	Effect of rationalization on flatfish fisheries	**Alan Haynie**
2:30	GOA rex sole	**Buck**

Tuesday June 12th

9:00	Gulf of Alaska Dover sole	**Buck**
11:00	Coffee break	
11:20	Gulf of Alaska Dover sole (continued)	**Buck**
12:30	Lunch	
1:30	Gulf of Alaska northern and southern rock sole	**Teresa**

Wednesday June 13th

9:00	Bering Sea yellowfin sole	**Tom and Jim**
11:00	Coffee break	
11:20	Bering Sea yellowfin sole (continued)	**Tom and Jim**
12:30	Lunch	
1:30	CIE panel discussion (assessment authors will be available)	