External Independent Peer Review by the Center for Independent Experts: SEDAR 38 South Atlantic and Gulf of Mexico King Mackerel Assessment Review

Miami, Florida 12-14th August 2014

Sven Kupschus

Prepared for

Center for Independent Experts (CIE)

Centre for Fisheries and Aquaculture Science
Lowestoft Laboratory
Pakefield Road
Lowestoft
Suffolk NR33 0HT
United Kingdom

Phone +44 1502 524454
e-mail: sven.kupschus@cefas.co.uk
www.cefas.co.uk

Content:

Executive Summary:

The SEDAR 38 Review Panel provided an independent peer review of key decisions and outputs from the Data and Assessment Workshops for South Atlantic and Gulf king. The review was held from 12-14 August 2014, in Miami, Florida. Data and assessment reports were presented to the Panel, and issues considered against the Review Panel's Terms of Reference (TORs) through open discussion. Additional analyses were requested at the review, and the results were considered. The Panel examined whether the Data and Assessment Workshop responses to their Terms of Reference were adequate, complete, and scientifically sound, and determined whether base-case analyses were preferred for determining stock status and developing management references.

The data used in the assessment models for both the South Atlantic and Gulf stocks were considered appropriate and scientifically sound. The major difference from previous assessments was the significant reduction in the size of the winter mixing zone and the reattribution of much of these catches from what was considered the Gulf stock to what is now considered the Atlantic stock. Unsurprisingly given the similarity of the data and the proximity of the stocks, there was a great deal of similarity in the assessment approaches for the two stocks, and in fact consistency between the models had been expressed as desirable.

Selectivity estimation for both models was relatively unconstrained, because previous assessments had indicated a significant dome-shaped selectivity for most fleets. However it was considered necessary to model at least one fleet as a more constrained logistic function. The choice of fleets for this differed between the models. In the South Atlantic tournament data, probably the most appropriate fleet to be considered truly logistic because of their targeting of the largest individuals, could quantitatively be ascribed a proportion of the total recreational catches. In the Gulf model no such separation was possible and instead the hand-line fishery males were considered to be selected asymptotically, with the female selectivity being treated as dome shaped. The review panel understandably had some concerns about the choice of dome-shaped selectivities due to the well known risk of developing cryptic biomass in such models, the presence or absence of which cannot be verified from data. However, based on some additional analysis, comparison with other models and estimation of the proportion of the biomass that was cryptic, it was determined that the choice of dome-shaped selectivities for the majority of the fleets was the most appropriate treatment of the data. Some improvements particularly with regards to model parsimony should be investigated.

Assessment models were set up to estimate the steepness of the stock recruitment relationship in order to describe the productivity of the stock. Both models converged on estimates (Gulf h=0.8, South Atlantic h=0.5) but the Gulf model required constraint in the form of a beta Bayesian prior. However neither stock recruit relationship was convincing in terms of its statistical properties and certainties as evaluated by the panel or the utility of the estimates in providing reference points. Consequently the panel recommended the application of a model identical to the aw-recommended model but without the estimation of a stock recruit relationship (recruitment independent of stock size). This improved the quality of estimates of current SSB, exploitation and recruitment, but complicated the evaluation of reference points and the appropriate recruitment estimates to take forward into the short-term projections. These issues were theoretically resolved at the review group, but there was insufficient time available to check whether the recommendations were procedurally consistent with what had been done for other stocks under similar circumstances, nor was it possible to carry out an exhaustive uncertainty analysis of the new models because of the

time required to complete the bootstrap analysis in SS3. The stock status determination given below is therefore conditional on the bootstrap analysis confirming the uncertainty estimation suggested by the likelihood information.

The best available information indicates that neither stock is in an overfished state, and that overfishing is not occurring and that further improvements to the assessment models which were suggested by the review panel were unlikely to materially alter this assessment. The panel concluded that the available data were appropriate and the models sufficiently robust to assess stock status against spawning potential based reference points and to conduct short term forecasts. The assessments indicated that biomass has improved in both stocks (since early 90's in the Gulf and since mid 90's on the Atlantic) in response to a period of above average recruitment. Exploitation of both stocks has declined in recent years, particularly dramatically in the Gulf stock, while in the Atlantic the decline has been more prolonged. In both cases exploitation levels are estimated to have reduced to levels not observed since the 60's during the development of the fisheries. Both assessments indicate that the last five recruitments have been considerably below average probably due to environmental / ecological effects, since they have come from the highest SSB levels in recent times.

TOR's 4 and 5 could not be addressed in their entirety and even for some of TOR3 the panel was limited to providing methodologies for the SSC to follow in order to reach the desired management advice, because not all the relevant results were available to the panel, mainly because the evaluation of the integrated model was complicated and took longer than scheduled. However without first completing TOR2 it was not possible to fully address subsequent TORs.

Background:

In accordance with the Statement of Work (SOW: Appendix 2), I was contracted to participate as a CIE independent review panelist for the 2014 SEDAR 38 Gulf and South Atlantic king mackerel review. This document represents my own findings and interpretation of the information provided, and is based on the panel meeting and discussions. However, some of the thoughts and conclusions were formulated in the process of writing this report, so may not be identical to or may go beyond those provided in the final official panel report.

Description of review activities:

This review was undertaken by Sven Kupschus at Cefas (Lowestoft, UK) and during the SEDAR Review Panel held in the Miami Florida, August 12-14th 2014. The stocks under consideration were the Gulf and South Atlantic king mackerel stocks.

The documentation (see bibliography, Appendix 1) was reviewed prior to the meeting. I actively participated in the SEDAR panel meeting in Miami and assisted with development of the SEDAR Review Panel meeting report. This separate report to CIE was completed on my return to Cefas.

The lead assessment scientists presented the individual data and assessment reports to the Panel, and issues were considered against the Review Panel's ToRs through open discussion.

In turn, additional sensitivity runs were requested by the Review Panel, including by myself, and further consideration of these results was made in Savannah. The Review Panel examined whether the Data and Assessment Workshop's responses to their TORs were adequate, complete, and scientifically sound, and determined whether the base-case analyses were appropriate for determining stock status and developing management references. Where there were small changes to the aw-recommended model that could be implemented in the available time this was done and TORs were evaluated against the RW-panel recommended model configuration.

Workshop results by term of reference:

1. **Evaluate the data used in the assessment, addressing the following:**

   a) Are data decisions made by the DW and AW sound and robust?

   In general the decisions carried forward from the data review process were inherently sensible and based on a good understanding of the fishery and the stock characteristics. It is difficult though to make these decisions a priori without an understanding of the modeled stock dynamics. In some instances the model treats the data unexpectedly so that the data and model development can often be an iterative process. Inconsistencies in the data derivations with the interpretation of the assessment model could be considered either under TOR1 or 2. Some issues are discussed under TOR1, others under TOR2 mainly for consistency with the way they have been treated in the panel report.

   The inclusion and exclusion of some datasets are described qualitatively and such decisions are difficult to verify from an external perspective. However the level of knowledge and understanding of the different data sources that are apparent from both the data and assessment workshop reports suggest that this has been conducted appropriately. The diligence with which duplication of data use within the model has been considered in the evaluation of both the indices and the age information is commendable.

   <u>Winter mixing Zone:</u>

   Since SEDAR 16, additional work on the stock definitions has improved the quantitative understanding of the spatial and temporal dynamics in the relative composition of the Atlantic and Gulf stocks in the area of mixing around the peninsula of Florida. Improvements in the understanding of the mixing dynamics in the western Gulf in relation to the sizeable Mexican fishery remain of a qualitative nature, because of the unavailability of Mexican landings information and thus carry a larger uncertainty. Though thought to be influential, the leaks in the population in the western Gulf were currently not considered sufficiently critical to invalidate status determination of the Gulf stock.

   In general I agree with these conclusions from the data workshop but have two comments that I think require further consideration.

   A change in the delineation of the mixing zone leading to a much reduced size of catches has led to a decrease in the uncertainty around the landings (now only 7%) that come from the mixed populations. I agree that the more detailed and up to date stock definition decreases the uncertainty in assessment status compared to previous results. However the mixing dynamics are still not understood. Past or future changes in the abundance of the respective stocks, or environmental effects on their distributions means that the true uncertainty is greater than the 7% quoted.

In addition, comments during panel discussions suggested that the reallocation of mixing zone catches was the major reasons for the more optimistic outlook in the assessment of the South Atlantic stock compared to previous assessments. In fact there is little trend over time in the landings from the mixing zone so that the reallocation of catches from the area merely serves to raise or lower the SSB trend and is hence likely to have little impact on the stock status estimation. The continuity VPA from SEDAR 16 altered by the addition of the new mixing zone catch allocation (RW03, Figure 36) still implies a more pessimistic outlook for the South Atlantic Stock. However the alteration of index weighting from the original setting to one based on inverse variance (and the removal of MRFSS data) provides a more consistent picture in SSB trend with that from the aw-recommended model. This similarity and the more objective treatment in SS3 of the index log-likelihoods suggests it is the change in assessment methodology, not the change in the mixing zone assumptions that has led to the more optimistic evaluation of the South Atlantic Stock.

Shrimp Discard data:

The discarding of king mackerel by the shrimp fisheries in both stocks is sufficiently significant to justify the estimation of the effect on the populations. This requires an absolute estimate of by-catch. CPUE of the shrimp fishery is assessed from observed commercial hauls and adjusted for the area and biomass effects by use of the SEAMAP trawl data. The adjusted CPUE is then multiplied by the commercial effort by area and year to derive absolute by-catches of king mackerel used in the assessment.

The SEAMAP trawl data are necessary in this analysis to fill in the gaps in the relatively sparse observer data. To avoid forcing the model to interpret the inter annual fluctuations in by-catch estimates as a recruitment signal SS3 is set up only to compare the mean observed value with the mean predicted value so that the two scale appropriately according to the shrimp effort, which is thought to be better known. I agree with the reasoning if shrimp by-catch provides no signal at all on recruitment. However, if the GLM analysis applied retains any interaction between the year effect and the fleet effect (i.e. differential distributions in the fleet between years) the shrimp by-catches will retain some signal of recruitment.

If this is the case the argument to me becomes circular. We trust the effort data more than the estimates of the discarded catches so that it has been implemented in the model as if known without error. By default the model then also has to accept the catch data as a perfect fit. These catches in the model are treated as having been removed from the population so in turn they do have a direct impact on the estimation of recruitment strength and subsequent length and age likelihood components, just not through the likelihood component of the index itself. I am not sure that I have a solution to the problem at hand, but I think investigations of the observer data in isolation from the SEAMPA trawl data could reveal if there is indeed a useful recruitment signal in the observer data, particularly as the assumed lack of such information does not preclude the effect on the current assessment methodology as suggested by the report.

Having said that, recent shrimp effort in the fleet is now so low and the widely implemented BRDs have reduced catches further so that the uncertainty associated with the by-catch in this fishery is unlikely to materially alter the conclusions of the assessment.

b) Are data uncertainties acknowledged, reported, and within normal or expected levels?

Uncertainties are appropriately acknowledged throughout the data workshop report. The majority of the uncertainties are based around biases, for example the suspected effect of alterations of bag limits on the MRFSS index, and are therefore not quantifiable outside of the model and only quantifiable inside the model if the appropriate dynamics can be identified and replicated. Some more information on the relative size of these uncertainties would be helpful in allowing reviewers to more closely follow the reasoning in model option choices. In my opinion the choices made were defensible and other appropriate choices would in my opinion not have lead to substantially different conclusions.

c) Are data applied properly within the assessment model?

This question is the reciprocal to TOR2 c and is addressed therein.

d) **Are input data series reliable and sufficient to support the assessment approach and findings?**

The review workshop found no major concerns with the development of the information going into either model. Some data series were excluded from the model on the basis that they suffered from significant biases or lack of independence with other data sources. This suggests a thorough investigation of the available data and a sound understanding of the modeling approach. Where comments in this CIE report question the data or its use in the context of the model, these refer to alternate options or ideas that could have been explored or applied. They do not represent definitive proof that such alternate treatments would have led to a more appropriate assessment. There was insufficient time at the workshop to examine alternatives in full detail and the effect on model parameterization of each of these individual options was judged to have been small. Consequently, I concur with the findings of the panel that by and large the input data were reliable and sufficient to support the assessment approach.

2. Evaluate the methods used to assess the stock, taking into account the available data.

a) Are methods scientifically sound and robust?

Both the Gulf and South Atlantic King Mackerel stocks were primarily assessed using SS3, but VPAs were also provided for continuity with previous assessments. SS3 was specifically designed to handle all possible data sources in a realistic and efficient manner. This flexibility is achieved through full integration and SS3's ability to maintain the uncertainty aspects of the data sources through the modeling process is its main strength. However it is also this flexibility that makes it difficult to understand the dynamics in the model. Invariably, the time taken to understand the information content of the data sources and how they interact with the modeled dynamics takes significantly longer than for simpler models. Consequently the identification of key uncertainties in relation to the data is much more difficult to evaluate. In contrast the key uncertainties with regards to the parameter estimates are much more easily identified than in many other models.

Here the VPA models provided for continuity provided a valuable short cut to determining some of the model inferences and some reassurance that SS3 was handling the data appropriately. The VPA models provided valuable insights into the major stock

dynamics such as selectivity and cohort strength and the implications of different data sources.

In summary, I concur with the panel that SS3 is an appropriate assessment tool for this stock (But also see comments in section 8).

b) Are assessment models configured properly and used consistent with standard practices?

The panel concluded that the assessment model was configured appropriately with regards to providing advice on stock status and projections. Some improvements in the specification of growth and selectivities were deemed possible.

Selectivities:

The immediate concern from the review panel with regards to both models was the flexibility with which selectivities were modeled. Standard practice should be to assume logistic selectivities, unless there are both statistical and mechanistic reasons to suggest otherwise. In both models selectivities were free to assume strongly dome shaped selectivities in length for most fleets and little evidence was provided at the review workshop that other options had been evaluated. Age selectivities were slightly less dome shaped, because of the increase in variability in size at older ages. The panels concerns were somewhat alleviated by the fact that at least at this time neither model had accumulated a large degree of cryptic biomass in the stocks (around 40% in each assessment) and the VPA at least for the Atlantic model suggested a similarly dome-shaped selectivity. Industry indicated at the review workshop that there were genuine financial incentives for selecting intermediate sized fish and that there were spatial, temporal and gear interactions by which they could separate different size categories of king mackerel.

I largely agree with the conclusions of the panel and do not believe that the following comments will substantially alter the outcome of these assessments, but feel that the approach taken here, if applied to other stocks some of which may be in more critical condition, will have a significant impact on management decisions.

Selectivities should by default be minimal in their flexibility, only when warranted by evidence of biased residuals or parameter estimates or dynamics that hit model bounds, should additional flexibility be allowed if mechanisms exist that could justify the chosen pattern of selectivity. Certainly I see little evidence for either model to be burdened with estimating the additional parameters required to develop separate selectivities for the sexes, as there is evidence neither that the sexes behave differently nor that the model estimates the parameters to be very different between sexes for the same fleet. When assessing stocks with sexual dimorphism using age selectivities such practices are necessary, but one of the major advantages of going to a length selectivity model is the reduction in parameters required.

The Gulf model:

The Gulf model did not have the benefit of being able to use the tournament fishery as a logistic fleet as in the South Atlantic. Instead, the handline fishery was trialed as a possibly logistic selectivity fleet. In the end only the males were retained in this form as it was argued that hand line caught females exhibited significant patterns in the length residuals. However the pattern of length residuals did not markedly improve by the relaxation, nor were they really much stronger than the length residual patterns in hand line caught males.

The argument that selectivity in at least one fishery must be logistic was not supported by the example of the Atlantic fishery, where a sensitivity analysis suggested there was little difference between the stock dynamics with a logistic or a dome-shaped selectivity for the tournament fleet. In any case, the Gulf model already contained a logistic fleet in terms of the plankton index. This was treated differently (as a biomass index) in the implementation of SS3 because it has no corresponding length distributions, but at least inherently it should act as a logistic selectivity function as the spawning biomass represents the integral of all mature ages.

The benefits of conducting further investigations may not be immediately apparent given the similarity of the outcome, but the approach, in principle at least, greatly increases the robustness of the assessments. The problem with over parameterised models is that they are able to ignore contrast or conflict in the information. When parameter estimates are independent this leads to convergence issues, but when parameters are partially correlated, as they frequently are in integrated models, this allows the model to 'hide' conflict in unexpected places.

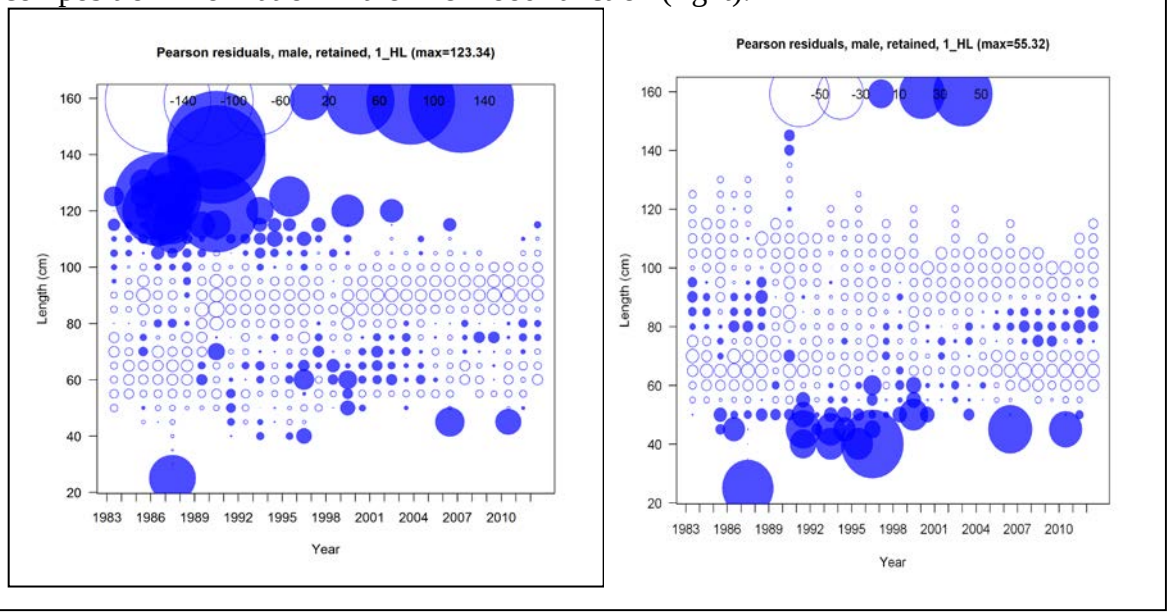
c) Are the methods appropriate for the available data?

As described under section 2a SS3 is specifically designed to deal with virtually all possible data sources and characteristics that regularly occur in fisheries data. Therefore it is not a question whether the method is appropriate for the available data, but is really more a question of whether the implementation is appropriate for the data.

Growth:

The review panel was concerned with the persistence of the length residual patterns in both models, particularly in light of the relative freedom that the models were given in treating selectivity. A significant amount of time was spent explaining this on the basis of the growth data, which did not inherently appear to fit VB growth, but the lack of fit could have been explained through selectivities which were estimated within the model. However the model appeared to treat the plus group data as 11 year-olds, rather than a

Figure 1: Differences in the length residual patterns for the male hand line catches between the base model (left) the same model without penalization of the age composition information in the likelihood function (right).



plus group, which meant it could not estimate the growth information appropriately to attain the observed lengths. This seemed to explain the apparent differences in stock dynamics between index and LF driven model and those including age information in the Gulf and the observed trends in stock dynamics when applying different weightings to the age and length data in the South Atlantic.

To confirm that the growth model rather than the selection pattern was the cause of the conflict I examined the length residuals of the hand line caught males of the base model and a sensitivity run using just the indices and the length information. Figure 1 shows this comparison and the change in residual pattern without a change in the shape of the selectivity strongly implicates the growth model as the cause. Ideally the growth model fitting should be corrected, but there was insufficient time to do this appropriately at the workshop. The addition of the age data did override much of the effect of the growth model parameterization so that it is unlikely that there is a major effect on the estimation of stock dynamics, though it cannot be excluded that some of the dome shape in the selectivity pattern arises from this problem.

3. Evaluate the assessment findings with respect to the following:

a) Are abundance, exploitation, and biomass estimates reliable, consistent with input data and population biological characteristics, and useful to support status inferences?

The review panel rejected the aw-recommended models based on the fact that both tried to estimate steepness within the model but that despite this there was very little evidence of a tangible stock recruitment curve in either model. Alternate models that did not implement stock recruit relationships were preferred by the panel (see section 3d). These rw-recommended models represent a fair representation of the stock dynamics of the stocks and as such I judge it to be an appropriate tool to provide stock status inferences. Not all quantities are equally well estimated through time, in particular the estimate of virgin recruitment is less certain than recent recruitments, while the opposite is true off the exploitation rates simply because landings were significantly smaller historically so the scale of F is much reduced. Biomass is estimated most accurately during the middle period, where cohorts are occluded, age information exists and the proportion of the stock in the cryptic biomass was lowest.

An additional comment on the model output was the metric of exploitation, the proportion of the stock (I assume in numbers rather than weight) captured each year. The problem with this is that the exploitation rate is fairly heavily influenced by the strong variability in recruitment in the stock. This variability will increase at lower stock sizes which in itself is not a problem if the interpretation of this value is consistent. I am unaware how the SSC interprets exploitation in its process of setting quotas. Certainly the information content in the ratio of exploitation/exploitation$_{MSY}$ is different from that one would derive from F/F$_{MSY}$ and rules used to set quotas would need to differ between these methods in order to attain the same level of risk. If this is implemented at the level of the SSC then the choice of exploitation metric is merely one of preference.

My personal view is that exploitation should as much as possible be reflective of the effects of fishing with effort being the ideal metric. Unfortunately the relationship between effort and mortality is often poorly defined and it is not possible to compare the effects of effort of different fleets because of the difference in selectivity. A compromise I favour is a F-based metric, the average F over the main exploitation ages (here 3-8).

b) Is the stock overfished? What information helps you reach this conclusion?

I concur with the review panel that neither of the stocks are overfished. Both assessment models indicate a positive recent response in SSB which is confirmed by a relative increase in the proportion of older ages in the data. The lowest observed SSB in both stocks is in the center of the period of lowest uncertainty on SSB and SSB has increased since then. Virgin SSB estimated is based on the assumption of average observed recruitment so assumes no dispensatory population growth so it also is conservative. The ratio of SSB_{lowest} / SSB_{virgin} is around 0.4 so that the current ratio must be above this value.

The review panel spent a disproportionate time within this TOR dealing with the consequences of the rejection of the steepness parameter from both models and the effects of this on the stock projections. The rejection of a discernible stock recruitment relationship meant that biomass reference points should be based on spawning potential ratios. However the group did not discuss what the appropriate ratio should be given the characteristics of the stocks.

The SPR_{30} had previously been implemented for this species so was considered the default. However, no investigations were made as to whether there was information on the stock dynamics to be gleaned from the new assessment that would suggest other options were more appropriate.

c) Is the stock undergoing overfishing? What information helps you reach this conclusion?

Quantitatively this conclusion is derived from the rw-recommended SS3 model output and the exploitation ratio. Because of recent low recruitment this ratio will be more conservative than an equivalent F-based metric (see comments in 3a) so it is highly likely that an equivalent F-based metric would provide the same conclusion with regards to stock status.

There was some discussion around fixing steepness at 0.8 and determining the equilibrium exploitation rate that would decrease the stock to 30% of virgin biomass as apparently conducted for other stocks under the gulf SSC umbrella. If this is a procedure that has been evaluated to be precautionary independent of any assumption about steepness and accepted by the SSC as such then this is an appropriate strategy. However the panel was unclear as to whether this was the case nor were we in a position to make that assessment in detail due to time constraints.

I feel that from a scientific perspective the approach is inappropriate. The SPR ratio is a way of judging precautionary exploitation when there is no evidence of a SR-relationship. The level of precaution is determined by what is known about the stock dynamics, in this case recruitment variability, natural mortality and growth. To then apply an additional level of precaution by implementing a steepness of 0.8 is inconsistent, since virgin recruitment is likely to be higher than observed recruitment at higher SSB levels so that the expectation of longterm yield is unrealistic. In addition it will have undesirable consequences on the projections, because in this case it suppresses estimates of current recruitment, which are in fact relatively well known. If the SSC is uncomfortable with the choice of SPR_{30} then they should simply chose a higher level of SPR and justify this in terms of their uncertainties rather than obscuring this in the procedures.

d) Is there an informative stock recruitment relationship? Is the stock recruitment curve reliable and useful for evaluation of productivity and future stock conditions?

This was discussed in great detail at the workshop and the panel concluded that there was no suitable evidence for a stock recruitment relationship in either stock. Both assessments indicated a reduced level of recruitment in recent years from increasing levels of biomass suggesting that recent recruitment at least was most heavily influenced by environmental or ecological conditions. Although in the South Atlantic the model statistically was able to estimate steepness the estimated value of 0.5 seemed low for the species, the confidence limits of the parameter estimate were unrealistically small and the retrospective analysis of the estimate suggested it was much more variable than suggested by the confidence limits. An external likelihood analysis, using the recruitment and SSB vectors suggested that the model estimates were inappropriate. The gulf model was unable to estimate steepness without a penalty on high values of steepness and the recruitment and SSB again suggested there was no plausible stock recruit relationship.

e) Are the quantitative estimates of the status determination criteria for this stock reliable? If not, are there other indicators that may be used to inform managers about stock trends and conditions?

I agree with the panel that the status determination criteria for the stock are reliable. The measures used are measures of exploitation and abundance as required, but any such measure carries with it some strengths and weaknesses in interpretation. The SSC should ensure that their interpretation is consistent with the measures as described in previous sections under TOR3.

4. Evaluate the stock projections, addressing the following:

a) Are the methods consistent with accepted practices and available data?

Fundamentally SS3 is designed to take forward the stock assessment estimates into the stock projections so the approach definitely falls under accepted practices when parameter estimates are accepted. What is more difficult to determine is how the different parameters within SS3 interact to provide the projections. In instances such as here where parameters are fixed, or knowledge on future recruitment is known.

In this case there were also discussions as to what recruitment to use, simple options included low, medium and high based on recruitment deviates. Auto correlated recruitment deviation vector estimates, which would increase the likelihood of a poor recruitment in the current situation (low recruitment for the last five years), or simple long term-average recruitment. There was no clear consensus because the cause of the recent recruitment reduction is not known beyond that it does not appear to be related to SSB. To complicate matters further the industry suggested that the most recent recruitment has been a good one based on their observations of small individuals, but there was as yet no data available either from the industry or the scientific survey to confirm the improvement in recruitment. There was insufficient time to investigate the risks involved in using different recruitment vector in relation to other possible outcomes. Consequently no attempt was made to reach a consensus on the appropriate recruitment vector to use in the projections.

Given the current state of the stock with SSB having increased and currently low exploitation rates compared to the past is seems unlikely that the choice is particularly critical in the instance of the Atlantic stock. This consideration is of course conditional on

the assumption that the bootstrap results that were not available to the panel are similar to hessian uncertainty estimates.

b) Are the methods appropriate for the assessment model and outputs?

Having rejected a SR-relationship in the South Atlantic within the model, stock status estimation by SPR and projections using average recruitment should have been straightforward. And would have been if one were to conduct these external to the model, but one of the benefits of using the integrated assessment model is that this can be done automatically.

Instead, it greatly complicated projections due to the integrated nature of SS3. Effectively turning off the stock recruit function by fixing steepness at 1.0 (This was methodological, not because we thought steepness should be 1.0. This was supposed to be a practical fix to attain the right input information for the forecast from the model. Scientifically it is only consistent with the commensurate application of an SPR-based biological reference point). However, without a stock recruit relationship in the model, the recruitment deviate vector started to be strongly constraining so that recruitments were unrealistically optimistic in recent years. The answer was to rerun the model again and but this time leave it to estimate the sigma r.

Due to time constraints and illness it was not possible to carry out similar preparatory work for the Gulf of Mexico stock, but it was discussed that the same approach should be applied, unless results differed qualitatively from those obtained for the South Atlantic through further within model parameter interactions. A cursory glance at the recruitment deviates from the rw-recommended model suggested to me that the same methodology should be applicable, but no information on the bias corrected recruitment trends was available to confirm this.

c) **Are the results informative and robust, and useful to support inferences of probable future conditions?**

I believe the stock projections are as robust as is possible in the case of the South Atlantic stock. It is however important to remember that the fishing mortalities required to drive the stock down to B_{SPR30} are higher than any F's that have been observed for any duration, nor has SSB ever been reduced to levels of 30% virgin biomass. Consequently, fishing at F_{SPR30} will eventually take the stock out of the previously observed range of stock dynamics. Although the extrapolation is relatively small, management should approach such areas of unknown stock dynamics knowingly and slowly.

d) Are key uncertainties acknowledged, discussed, and reflected in the projection results?

No uncertainty estimates of stock projections were possible. The only projection results that were available to the panel for review were those for yield in the South Atlantic stock using the hessian standard error.

A key uncertainty that should be investigated beyond the usual effect of the assessment parameterization is the likely impact of future recruitment choices as detailed earlier in this section.

5. Consider how uncertainties in the assessment, and their potential consequences, are addressed.

The advantage of SS3 is that it integrates the key uncertainties in the estimated parameters of the assessment and carries the appropriate risks through to the stock status

estimation and projections. Therefore it provides a one-stop-shop for all uncertainties within the model. Theoretically the hessian estimates of the uncertainties in the parameters describe the uncertainty in any of the output parameters and therefore address the requirements of this TOR with regards to the quantities estimated in the models. However the output in uncertainties for SS3 models is usually based on bootstrap simulations which were not available at the time of the review due to the panel decision to reject the estimate of steepness from both models. The results of the marginal probabilities for a specific parameter are usually very similar to the hessian results, but the joint probabilities can be strongly asymmetric so that the panel felt uncomfortable in making a full assessment of the uncertainty criteria in the absence of the bootstrap results.

However process error and fixed parameter estimates cannot be addressed within the model and these need to be considered independently though sensitivity analyses. In this case there are three uncertainties that need to be considered:

Growth:

There is a misspecification in the growth model. The models using the age data minimizes the effects on the estimates of number-at-age in the population, however if in projections the proportion of fish greater than age 10 increases, SSB and to a lesser degree yield estimates (because of the dome-shaped selectivities) will be under estimated. In contrast, if younger fish increase disproportionately in the population, then SSB is likely to be overestimated. Current estimates of SSB are scaled by the catches so carry only minimal bias.

Natural mortality:

Natural mortality is estimated external to the model. Such estimations are notoriously difficult and usually thought not to be time invariant. The effect of the uncertainty in these assessments is unusually large mainly because natural mortality represents a much larger proportion of the total mortality because the stock is so lightly exploited. This has relatively little effect in the SSB/SSB_{MSY} or F/F_{MSY} ratios of the South Atlantic base assessment during the data rich period (where indices, length and age information are available). However this model was able to compensate by changing the estimate of *h* whereas this has been removed in the rw-recommended model. Consequently it is important to examine the effects on stock status estimation for the new South Atlantic assessment.

The base model for the Gulf appeared more dramatic, with reductions of M-0.1 to M+0.1 spanning the gamut of SSB_{cur} being well below SSB_{MSY} to near virgin SSB. Certainly the lower end of the sensitivity range explored seem unrealistic, so there is unlikely to be a change in the determination of stock status in the base model from a realistic range of M's alone. However as in the South Atlantic the SR relationship has been removed and the inability of the model to now respond to different Ms by altering productivity is greatly curtailed so the effects may be more significant for the rw-recommended model.

Productivity:

As mentioned, the rw-recommended model rejected any possible stock recruitment relationships in both of the assessments that significantly impacted recruitment size at observed biomass levels. However this does not exclude the possibility that at lower levels of biomass such a relationship might become apparent. This uncertainty is acknowledged in the fact that the panel recommended that the stock be managed to an appropriate SPR value with a commensurate F value that would not rapidly take the stock out of the range of observed SSB values.

Future recruitment:

Several options for likely future recruitment are considered under the projection settings. These can be implemented in SS3 to create stock projections but the uncertainty in the future level of recruitment has not been resolved, see the section 4a.

Other:

A list of other uncertainties such as the maturity ogive which has changed substantially since the previous assessment, the amount of mixing between the stocks in relation to where the catches are taken were also discussed but I consider the risks originating from these external sources of uncertainty small compared to those listed above.

6. Consider the research recommendations provided by the Data and Assessment workshops and make any additional recommendations or prioritizations warranted.

- Provide recommendations on possible ways to improve the SEDAR process.

SS3 is an excellent research tool with a vast number of different implementation options capable of representing an almost infinite array of stock and fleet dynamics. But it is extremely time consuming, both in the development of assessments and in the evaluation / review. The short time now available to reviewers and the assessment panel to provide additional requests does hamper the ability to gain a complete understanding of all the implemented dynamics as SS3 is being applied more and more in US led stock assessments. In the case of SEDAR 38 the implementation of the two models was very similar so that inference on the understanding gained on one model was ultimately transferable to the other model. Despite this a disproportionate amount of time was still devoted to the evaluation of the model and management reference point decisions and forecasts were considered only in relative brevity. Because of the changes requested to the base model no bootstrap run was available for either stock by the end of the workshop, so that panel inference on the uncertainty was based only on the hessian approximation. From my experience this is usually a reasonable assumption, but the conclusions reached here are conditional on there being good correspondence between the two.

Although a number of TORs were not addressed in their entirety, I feel in the end the panel managed to address most to the point where enough guidance is provided to the SSC in their decision making process and choices between the remaining options are more a matter of policy and / or convention in acceptable levels of risk. The exception here is definitely TOR 6 and to a lesser degree TOR5. In general the panel and the analysis worked together well so it is unfortunate that we ran out of time to address all TORs fully.

In some ways I think SS3 is in part to be blamed for this in the sense that it offers such levels of complexity that it is easy to over parameterise models, which then exhibit unexpected properties due to implicit parameter correlations. On the other hand, while SS3 offers these options there is no requirement of assessors to implement these options and more simple model setups are certainly possible and in my opinion desirable. For example In SEDAR 38 male and female selectivity-at-length was modeled independently for most fleets mainly because it was possible, but it was not necessary (selectivities for sexes were very similar between fleets). Given the known problems in how female selectivity is modeled as an offset to male selectivity in SS3 this represents in my opinion an unnecessary risk of either rejecting the model because its properties were not understood or providing advice on the basis of a flawed model because the risks were not

understood. No doubt due to my excellent co-reviewers and the diligent assessment panel this does not apply in this instance but certainly the other aspects of the review, such as the evaluation of the uncertainty and forecast methods, did experience less detailed reviews compared to the assessment itself and leave some work to be concluded after the review.

I think that model parsimony should be given more consideration in development of assessment models and the SEDAR process needs to consider if the use of more complex models is compatible with their requirements for risk evaluation by review panels. This needs to consider that the more complex a model is the higher the risk that some misspecifications go unnoticed or the greater the effort required to reach the same level of quality control. There is a tradeoff here, and guidance to assessors and reviewers by the respective SSCs would be welcome if not necessary.

7. **Provide guidance on key improvements in data or modeling approaches which should be considered when scheduling the next assessment.**

The implementation of the growth function needs to be changed. This is largely a methodological issue on whether the input information is reformatted to go to older ages, or whether it is more efficient to use the implementation to estimate L10 rather than Linf. Increasing the number of age groups in the model is a little more work, but to me seems the more sensible option in the long run, particularly as some more information on these older ages will help the model better define the descending limb of the selectivities. Proper parameterization of growth is likely to reduce the dome-shape in the selectivities although it seems unlikely that it will remove it entirely for all fleets.

In general model parsimony needs a rethink. Usually the decision on whether to add additional parameters is largely based on whether the model continues to converge. Even when based on the AIC, a parameter is statistically significant it does not necessarily imply that it is sensible, or even helpful for the assessment as a whole. Saturating models with parameters generally makes them sensitive to small changes in either data or assumptions, which reduces the robustness in the advice and complicates model evaluation.

Despite the fact that I think these models are sufficiently robust to be used for management advice, I still think they are over parameterized and in the indications were that stock status were more critical I would almost certainly have felt uncomfortable in making that statement.

Specifically in these models the parameters used to estimate selectivity can be reigned in significantly, definitely in terms of the gender specific selectivity at length, but I feel also in the estimation of the descending limb for some of the selectivity functions. This should include determination of whether the SEMAP plankton index does not serve as a logistic selectivity index in Gulf model (it should) which should then in turn allow for using the same length selectivity function for the hand line fleet.

I would like to see an independent analysis (without the survey information) of the shrimp by-catch data to determine if it does contain a suitable index of recruitment that could be used to provide additional information to the SEAMAP trawl survey index. As used at present in the model any recruitment signal remaining in the shrimp data will fight the signal from the survey even if the two are the same. This should be avoided.

APPENDIX 1: Bibliography of documents provided to the review panel:

SEDAR 38 South Atlantic and Gulf of Mexico King Mackerel Document List SEDAR 38 Document List 10-Jan-14

Document #	Title	Authors	Date Submitted
Documents Prepared for the Data Workshop			
SEDAR38-DW-01	King mackerel (*Scomberomorus cavalla*) larval indices of relative abundance from SEAMAP Fall Plankton Surveys, 1986 to 2012	David S. Hanisko and Joanne Lyczkowski-Shultz	10 Dec 2013
SEDAR38-DW-02	King mackerel abundance indices from SEAMAP groundfish surveys in the Northern Gulf of Mexico	Adam G. Pollack and G. Walter Ingram, Jr.	10 Dec 2013 Addendum – 30 Dec 2013
SEDAR38-DW-03	King mackerel abundance indices from NMFS small pelagics trawl surveys in the Northern Gulf of Mexico	Adam Pollack and G. Walter Ingram, Jr.	10 Dec 2013
SEDAR38-DW-04	Standardized catch indices of king mackerel from the U.S. Marine Recreational Fisheries Statistics Survey, 1981 to 2012	Matthew Lauretta and John F. Walter	22 Nov 2013
SEDAR38-DW-05	SEDAR standardized report cards used for review of indices of abundance for Atlantic and Gulf of Mexico king mackerel	SEDAR 38 Indices Working Group	7 January 2014

SEDAR38-DW-06	Standardized catch rates of Atlantic king mackerel (*Scomberomorus cavalla*) from the North Carolina Commercial fisheries trip tickets 1994-2013	John Walter and Stephanie McInerny	22 Nov 2013
SEDAR38-DW-07	Analysis of environmental factors affecting king mackerel landings along the east coast of Florida	Peter J. Barile	22 Nov 2013
SEDAR38-DW-08	Analysis of annual, monthly and weekly king mackerel landings in the east FL "mixing zone" : evidence of stock migrations and a "resident" population on the east coast of FL	Peter J. Barile	22 Nov 2013
SEDAR38-DW-09	Sampling History of the King Mackerel Commercial Fisheries in the Southeastern United States by the Federal Trip Interview Program (TIP)	Courtney R. Saari	22 Nov 2013
SEDAR38-DW-10	Standardized catch rates of from commercial logbook data for king mackerel from the United States Gulf of Mexico, South Atlantic, and Mixing Zone, 1993-2013	John F. Walter and Kevin J. McCarthy	6 January 2014
SEDAR38-DW-11	King mackerel index of abundance in coastal US South Atlantic waters based on a fishery-independent trawl survey	Tracey I. Smart and Jeanne Boylan	22 Nov 2013 Addendum – 30 Dec 2013

SEDAR38-DW-12	Trends from Non-CPUE Standardized King mackerel Landing Logs from Long Bay, South Carolina Recreational Pier Fishery	Christian Johnson	22 Nov 2013
SEDAR38-DW-13	King Mackerel Historical Pictures Summary	Rusty Hudson	22 Nov 2013
SEDAR38-DW-14	SEDAR 16 King Mackerel Review Panel Information Provided by Ben Hartig	Ben Hartig	29 Nov 2013
SEDAR38-DW-15	A review of Gulf of Mexico and Atlantic king mackerel *(Scomberomorus cavalla)* age data, 1986 – 2013, from the Panama City Laboratory, Southeast Fisheries Science Center, NOAA Fisheries Service	Chris Palmer, Doug DeVries, Carrie Fioramonti, and Hannah Lang	3 Dec 2013 Addendum: 7 January 2014
SEDAR38-DW-16	Updated standardized catch rates of king mackerel (*Scomberomorus cavalla*) from the headboat fishery in the U.S. Gulf of Mexico and U.S. South Atlantic	Matt Lauretta and Shannon L. Cass-Calay	6 Dec 2013
SEDAR38-DW-17	Historical For-Hire Fishing Vessels South Atlantic Fishery Management Council 1930s to 1985	Rusty Hudson	3 January 2014
SEDAR38-DW-18	Historical photographs of For-Hire Fishing Vessels 1930s to 1985	Rusty Hudson	3 January 2014

Documents Prepared for the Assessment Process

SEDAR38-AW-01	Growth models for king mackerel from the south Atlantic and Gulf of Mexico	Linda Lomdardi	9 May 2014
SEDAR38-AW-02	Addendum to "SEDAR 38-10": New South Atlantic logbook index based upon revised mixing zone definition and new indices for the Gulf and South Atlantic using only trolling gear	John Walter	10 March 2014
SEDAR38-AW-03	The NMFS-SEFSC must account for climate change and inter-annual environmental variability in all South Atlantic stock assessments	Peter J. Barile	10 March 2014
SEDAR38-AW-04	Can climate explain temporal trends in king mackerel (Scomberomorus cavalla) catch-per-unit-effort and landings?	Harford,W.J, Sagarese,S.R., Nuttall,M.A., Karnauskas,M., Liu,H., Lauretta,M., Schirripa,M. & Walter,J.F.	14 July 2014
SEDAR38-AW-05	Age frequency distributions, age length keys, length at ages, and sex ratios for king mackerels in the Gulf of Mexico and South Atlantic from 1986-2013	Ching-Ping Chih	20 March 2014
SEDAR38-AW-06	Length frequency distributions for king mackerels in the Gulf of Mexico and South Atlantics from 1978-2013	Ching-Ping Chih	20 March 2014

Documents Prepared for the Review Workshop

SEDAR38-RW-01	South Atlantic Shrimp fishery bycatch of king mackerel	John Walter and Jeff Isley	6 August 2014
SEDAR38-RW-02	Methods Used to Compile South Atlantic Shrimp Effort Used in the Estimation of King Mackerel Bycatch in the South Atlantic Shrimp Fishery	David Gloeckner	5 August 2014
SEDAR38-RW-03	Virtual population analysis for Atlantic king mackerel	Matthew Lauretta	4 August 2014
SEDAR38-RW-04	Virtual population analysis of Gulf of Mexico king mackerel	Matthew Lauretta	4 August 2014
SEDAR38-RW-05	King Mackerel and Spanish Mackerel larval data on the northeast U.S. Shelf	Harvey J. Walsh, David E. Richardson, Katrin E. Marancik, and Jon A. Hare	22 July 2014

Final Stock Assessment Reports

SEDAR38-SAR1	King mackerel: Gulf of Mexico Migratory Group	SEDAR 38 Panels
SEDAR38-SAR2	King mackerel: South Atlantic Migratory Group	SEDAR 38 Panels

Reference Documents

SEDAR38-RD01	Spatial and temporal variability in the relative contribution of king mackerel (*Scomberomorus cavalla*) stocks to winter mixed fisheries off South Florida	Todd R. Clardy, William F. Patterson III, Douglas A. DeVries, and Christopher Palmer

SEDAR38-RD02	King mackerel population dynamics and stock mixing in the United States Atlantic Ocean and Gulf of Mexico	Katherine E. Shepard
SEDAR38-RD03	A Cooperative Research Approach to Estimating Atlantic and Gulf of Mexico King Mackerel Stock Mixing and Population Dynamics Parameters	William F. Patterson III and Katherine E. Shepard
SEDAR38-RD04	Contemporary versus historical estimates of king mackerel (*Scomberomorus cavalla*) age and growth in the U.S. Atlantic Ocean and Gulf of Mexico	Katherine E. Shepard, William F. Patterson III, Douglas A. DeVries, and Mauricio Ortiz
SEDAR38-RD05	Trends in Atlantic contribution to mixed-stock king mackerel landings in South Florida inferred from otolith shape analysis	Katherine E. Shepard, William F. Patterson III, and Douglas A. DeVries
SEDAR38-RD06	Coastal upwelling in the South Atlantic Bight: A revisit of the 2003 cold event using long term observations and model hindcast solutions	Kyung Hoon Hyun and Ruoying He
SEDAR38-RD07	FishSmart: An Innovative Role for Science in Stakeholder-Centered Approaches to Fisheries Management	Thomas J. Miller , Jeff A. Blair , Thomas F. Ihde , Robert M. Jones, David H. Secor & Michael J. Wilberg
SEDAR38-RD08	FishSmart: Harnessing the Knowledge of Stakeholders to Enhance U.S. Marine Recreational Fisheries with Application to the Atlantic King Mackerel Fishery	Thomas F. Ihde, Michael J. Wilberg, David H. Secor, and Thomas J. Miller
SEDAR38-RD09	SEDAR 16 Final Document List	SEDAR 16 Panels
SEDAR38-RD10	History of fishing in Ponce Inlet	The Quarterly Newsletter of the Ponce de Leon Inlet Lighthouse Preservation Association, Inc.

SEDAR38-RD11	Biological-Statistical Census of the Species Entering Fisheries in the Cape Canaveral Area	William W. Anderson and Jack W. Gehringer

Appendix 2 Statement of work:

Attachment A: Statement of Work for Dr. Sven Kupschus (CEFAS)

External Independent Peer Review by the Center for Independent Experts

SEDAR 38 South Atlantic and Gulf of Mexico King Mackerel Assessment Review

Scope of Work and CIE Process: The National Marine Fisheries Service's (NMFS) Office of Science and Technology coordinates and manages a contract providing external expertise through the Center for Independent Experts (CIE) to conduct independent peer reviews of NMFS scientific projects. The Statement of Work (SoW) described herein was established by the NMFS Project Contact and Contracting Officer's Representative (COR), 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: SEDAR 38 will be a compilation of data, benchmark assessments of the stocks, and an assessment review conducted for South Atlantic and Gulf of Mexico king mackerel. The review panel is ultimately responsible for ensuring that the best possible assessments are provided through the SEDAR process. The stocks assessed through SEDAR 38 are within the jurisdiction of the South Atlantic and Gulf of Mexico Fishery Management Councils, and the state waters of Texas, Louisiana, Mississippi, Alabama, Florida, Georgia, South Carolina, and North Carolina. The Terms of Reference (ToRs) of the peer review are attached in **Annex 2**.

Requirements for CIE Reviewers: Three CIE reviewers shall have the necessary qualifications to complete an impartial and independent peer review in accordance with the tasks and ToRs described in the SoW herein. The CIE reviewers shall have expertise in stock assessment, statistics, fisheries science, and marine biology sufficient to complete the tasks of the scientific peer-review described herein. 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 participate and conduct an independent peer review during the SEDAR 33 panel review meeting scheduled in Miami, Florida during August 12-14, 2014.

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

Tasks prior to the meeting: The contractor shall independently select qualified reviewers that do not have conflicts of interest to conduct an independent scientific peer review in

accordance with the tasks and ToRs within the SoW. Upon completion of the independent reviewer selection by the contractor's technical team, the contractor shall provide the reviewer information (full name, title, affiliation, country, address, email, and FAX number) to the contractor officer's representative (COR), who will forward this information to the NMFS Project Contact no later than the date specified in the Schedule of Milestones and Deliverables. The contractor shall be responsible for providing the SoW and stock assessment ToRs to each reviewer. The NMFS Project Contact will be responsible for providing the reviewers with the background documents, reports, foreign national security clearance, and other information concerning pertinent meeting arrangements. The NMFS Project Contact will also be 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 COR prior to the commencement of the peer review.

Foreign National Security Clearance: The reviewers shall participate during a panel review meeting at a government facility, and the NMFS Project Contact will be responsible for obtaining the Foreign National Security Clearance approval for the reviewers who are non-US citizens. For this reason, the reviewers shall provide by FAX (not by email) the 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/.

Pre-review Background Documents: Approximately two weeks before the peer review, the NMFS Project Contact will send (by electronic mail or make available at an FTP site) to the COR the necessary background information and reports (i.e., working papers) for the reviewers to conduct the peer review, and the COR will forward these to the contractor. In the case where the documents need to be mailed, the NMFS Project Contact will consult with the COR on where to send documents. The reviewers are responsible only for the pre-review documents that are delivered to the contractor in accordance to the SoW scheduled deadlines specified herein. The reviewers shall read all documents deemed as necessary in preparation for the peer review.

Tasks during the panel review meeting: Each reviewer shall conduct the independent peer review in accordance with the SoW and stock assessment ToRs, and shall not serve in any other role unless specified herein. **Modifications to the SoW and ToRs shall not be made during the peer review, and any SoW or ToRs modifications prior to the peer review shall be approved by the COR and contractor.** Each 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 stock assessment ToRs as specified herein. The NMFS Project Contact will be responsible for any facility arrangements (e.g., conference room for panel review meetings or teleconference arrangements). The NMFS Project Contact will also be responsible for ensuring that the Chair understands the contractual role of the reviewers as specified herein. The contractor can contact the COR and NMFS Project Contact to confirm any peer review arrangements, including the meeting facility arrangements.

Tasks after the panel review meeting: Each reviewer shall prepare an independent peer review report, and the report shall be formatted as described in **Annex 1**. This report should

explain whether each stock assessment ToR was or was not completed successfully during the SEDAR meeting. If any existing BRP or their proxies are considered inappropriate, each independent report shall include recommendations and justification for suitable alternatives. If such alternatives cannot be identified, then the report shall indicate that the existing BRPs are the best available at this time. Additional questions and pertinent information related to the assessment review addressed during the meetings that were not in the ToRs may be included in a separate section at the end of an independent peer review report.

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

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

1) Conduct necessary pre-review preparations, including the review of background material and reports provided by the NMFS Project Contact in advance of the peer review.
2) Participate during the panel review meeting at Miami, Florida during August 12-14, 2014.
3) Conduct an independent peer review in accordance with the ToRs (**Annex 2**).
4) No later than August 25, 2014, each CIE reviewer shall submit an independent peer review report addressed to the "Center for Independent Experts," and sent to Mr. Manoj Shivlani, CIE Lead Coordinator, via email to shivlanim@bellsouth.net, and CIE Regional Coordinator, via email to Dr. David Sampson at david.sampson@oregonstate.edu. Each CIE report shall be written using the format and content requirements specified in Annex 1, and address each ToR in **Annex 2**.

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

6 July 2014	CIE sends reviewer contact information to the COR, who then sends this to the NMFS Project Contact
29 July 2014	NMFS Project Contact sends the stock assessment report and background documents to the CIE reviewers.
12-14 August 2014	Each reviewer shall conduct an independent peer review during the panel review meeting in Miami, Florida
25 August 2014	CIE reviewers submit draft CIE independent peer review reports to the CIE Lead Coordinator and CIE Regional Coordinator
8 September 2014	CIE submits CIE independent peer review reports to the COR
15 September 2014	The COR 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 COR within 10 working days after receipt of all required information of the decision on changes. The COR 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 COR 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 COR (William Michaels, via William.Michaels@noaa.gov).

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

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

Support Personnel:

William Michaels, Program Manager, COR
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

Allen Shimada
NMFS Office of Science and Technology
1315 East West Hwy, SSMC3, F/ST4, Silver Spring, MD 20910
Allen.Shimada@noaa.gov Phone: 301-427-8174

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:

NMFS Project Contact:

Julie A. Neer, SEDAR Coordinator
4055 Faber Place Drive, Suite 201
North Charleston, South Carolina 29405
Julie.Neer@safmc.net Phone: 843-571-4366

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. The CIE independent report shall be a stand-alone document for others to understand the weaknesses and strengths of the science reviewed. The CIE independent report shall be an independent peer review of each ToRs.

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

Annex 2: Tentative Terms of Reference for the Peer Review

SEDAR 38 South Atlantic and Gulf of Mexico King Mackerel Assessment Review

1. Evaluate the data used in the assessment, addressing the following:
   e) Are data decisions made by the DW and AW sound and robust?
   f) Are data uncertainties acknowledged, reported, and within normal or expected levels?
   g) Are data applied properly within the assessment model?
   h) Are input data series reliable and sufficient to support the assessment approach and findings?
2. Evaluate the methods used to assess the stock, taking into account the available data.
   d) Are methods scientifically sound and robust?
   e) Are assessment models configured properly and used consistent with standard practices?
   f) Are the methods appropriate for the available data?
3. Evaluate the assessment findings with respect to the following:
   f) Are abundance, exploitation, and biomass estimates reliable, consistent with input data and population biological characteristics, and useful to support status inferences?
   g) Is the stock overfished? What information helps you reach this conclusion?
   h) Is the stock undergoing overfishing? What information helps you reach this conclusion?
   i) Is there an informative stock recruitment relationship? Is the stock recruitment curve reliable and useful for evaluation of productivity and future stock conditions?
   j) Are the quantitative estimates of the status determination criteria for this stock reliable? If not, are there other indicators that may be used to inform managers about stock trends and conditions?
4. Evaluate the stock projections, addressing the following:
   e) Are the methods consistent with accepted practices and available data?
   f) Are the methods appropriate for the assessment model and outputs?
   g) Are the results informative and robust, and useful to support inferences of probable future conditions?
   h) Are key uncertainties acknowledged, discussed, and reflected in the projection results?
5. Consider how uncertainties in the assessment, and their potential consequences, are addressed.
   - Comment on the degree to which methods used to evaluate uncertainty reflect and capture the significant sources of uncertainty in the population, data sources, and assessment methods.

- Ensure that the implications of uncertainty in technical conclusions are clearly stated.

6. Consider the research recommendations provided by the Data and Assessment workshops and make any additional recommendations or prioritizations warranted.
   - Clearly denote research and monitoring that could improve the reliability of, and information provided by, future assessments.
   - Provide recommendations on possible ways to improve the SEDAR process.
7. Provide guidance on key improvements in data or modeling approaches which should be considered when scheduling the next assessment.
8. Prepare a Peer Review Summary summarizing the Panel's evaluation of the stock assessment and addressing each Term of Reference.
   - 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.

Annex 3: Tentative Agenda for

SEDAR 38 South Atlantic and Gulf of Mexico King Mackerel Assessment Review

Tentative Agenda

Miami, Florida
12-14 August 2014

Tuesday

9:00 a.m.	**Introductions and Opening Remarks**	
	Coordinator	
	- Agenda Review, TOR, Task Assignments	
9:30 a.m. – 11:30 a.m.	**Assessment Presentations**	**TBD**
11:30 a.m. – 1:00 p.m.	**Lunch Break**	
1:00 p.m. – 6:00 p.m.	**Continue Presentations/ Panel Discussion**	**Chair**
	- Assessment Data & Methods	
	- Identify additional analyses, sensitivities, corrections	

***Tuesday Goals*:** Initial presentations completed, sensitivity and base model discussion begun

Wednesday

8:00 a.m. – 11:30 a.m.	**Panel Discussion**	**Chair**
	- Assessment Data & Methods	
	- Identify additional analyses, sensitivities, corrections	
11:30 a.m. – 1:00 p.m.	**Lunch Break**	
1:00 p.m. – 6:00 p.m.	**Panel Discussion/Panel Work Session**	**Chair**
	- Continue deliberations	
	- Review additional analyses	
	- Recommendations and comments	

***Wednesday Goals*:** sensitivities and modifications identified, preferred models selected, projection approaches approved, Report drafts begun

Thursday

8:00 a.m. – 11:30 a.m.	**Panel Discussion**	**Chair**
	- Final sensitivities reviewed.	
	- Projections reviewed.	**Chair**
11:30 a.m. – 1:00 p.m.	**Lunch Break**	
1:00 p.m. – 5:00 p.m.	**Panel Discussion or Work Session**	**Chair**
	- Review Reports	
5:00 p.m.	**ADJOURN**	

Thursday Goals: Complete assessment work and discussions, final results available. Draft Reports reviewed.