Center for Independent Experts (CIE)

Independent Peer Review Report of the Red Hake Stock Structure Research Track

Conducted by: Haritz Arrizabalaga

April 2020

Northeast Fisheries Science Center, Woods Hole, MA

Table of Contents

Executive Summary

The Red Hake Stock Structure Research Track Stock Assessment Review Committee (SARC) met at NOAA's Northeast Fisheries Science Center in Woods Hole, MA during March 9th – 12th, 2020. The terms of reference (ToR) were diverse and related to the definition of stock structure and evaluation of stock status of red hake, as well as prioritization of future research. Accordingly, the Red Hake Stock Structure Working Group (WG) incorporated a wide range of expertise to accommodate the multidisciplinary nature of the tasks. The WG conducted very high-quality work and all ToRs were successfully met, although additional sensitivity tests are needed on the Reference Points estimated under ToR 5 before they are used for management advice.

The WG did an excellent job in compiling available information and conducted several analyses to enlighten potential red hake stock structure, including fisheries data, scientific trawl survey data, age data, otolith chemistry data as well as spawning, larval and young-of-the-year (YOY) distribution data. They determined that there was not sufficient evidence to reject the null hypothesis of two stocks, a northern stock that encompasses the northern Georges Bank and the Gulf of Maine and a southern stock that encompasses the southern New England and the southern Georges Bank. The group properly identified the analyses that were supporting this as well as alternative hypotheses. The two-stock hypothesis was mostly supported by phenotypic differences (in growth, meristics and otolith morphology) as well as the consistency of relative abundance indices in those two areas as determined by the "Management Unit Estimator", a statistical analysis designed to define stock boundaries. On the other hand, the single stock hypothesis was mostly supported by spawning and early life history analyses, with some additional support from the otolith chemistry study.

I concur with the WG recommendation to maintain the current two-stock separation for red hake. However, this decision largely relies on the fact that there is no clear better alternative. Important uncertainties around red hake stock structure remain, which might have important implications on the success of the U.S. stock assessment and management process, thus further research is recommended in the future to try to improve the situation.

The analyses of the survey catchability data where very comprehensive, well designed and conducted with state-of-the-art statistical methods, following previous experience for flatfish. I agree that the catchability estimates can now be used in the assessment. More so when the WG got comparable estimates in a preliminary analysis with an independent dataset based on the HABCAM survey.

The study on the effect of the net wing spread on catchability was also very comprehensive. Results suggest that, contrary to expectations, gear efficiency does not decrease when the net is overspread or underspread. Anyway, the WG correctly recommended that future calculations of swept area biomass should explore directly accounting for measured wingspread rather than applying the average wingspread to all tows.

On ToR5, the WG applied the existing model framework (An Index Model, AIM) to the northern and southern stocks. Both models were not significant, which suggests that, under the assumed stock structure, fishing mortality is not driving abundance. The WG correctly determined that AIM should not be used in future stock assessments and considered spawning potential ratio reference points as an alternative.

The WG proposed $F_{40\%}$ and $SSB_{40\%}$ as proxies for F_{MSY} and SSB_{MSY} respectively, which are reasonable and widely used for a range of stocks. On this basis, I think that these reference

points can be useful in future stock assessment and management applications. However, I recommend that the particular $F_{40\%}$ and $SSB_{40\%}$ estimates for northern and southern red hake should not be directly used to set catch advice in the short term until further sensitivity analyses are conducted with respect to some of the assumptions, such as natural mortality, selectivity and recruitment levels. The WG assumed a single value of M, but there are some observations that suggest that M might be changing. The WG also used recent (2009-2019) recruitment values for estimating $SSB_{40\%}$, but there are signs of productivity shifts that might affect the reference point estimates.

In spite of the concerns around the reference points, exploitation rates for both stocks are very low and overfishing is likely not occurring. Survey data also indicate that the biomass in northern areas is increasing, thus it is also likely that the northern stock is not overfished. Conversely, the dynamics of the southern stock are quite different and the status of the stock is more uncertain. Survey data indicate that red hake in southern areas has decreased very substantially and remains at very low values since the 1990s, despite low catch levels (compared to historical highs). A decrease in stock productivity might have occurred and exploitation might not be driving biomass trends if natural mortality is much higher than fishing mortality. If this is the case, it might not be possible to return to historical productivity levels by just further reducing exploitation on southern red hake. Thus, although I recommend performing sensitivity analyses of $SSB_{40\%}$ with respect to higher equilibrium recruitment levels, it might not be reasonable to consider such reference points as a target.

Last, in ToR6, the WG identified the main knowledge gaps and developed a prioritized list of recommendations to address them. I generally agree with the comprehensive set of recommendations provided by the group, especially on the need for the genetic study. If possible, I would recommend conducting the genetic study coupled with the otolith chemistry study and the analysis of other natural tags, as they complement each other. I also recommend having full spatial coverage during the ichthyoplankton surveys as they are of utmost importance to identify potential spawning areas.

Background

The National Marine Fisheries Service (NMFS) is mandated by the Magnuson-Stevens Fishery Conservation and Management Act, Endangered Species Act, and Marine Mammal Protection Act to conserve, protect, and manage national marine living resources based upon the best scientific information available. NMFS science products, including scientific advice, are often controversial and may require timely scientific peer reviews that are strictly independent of all outside influences. A formal external process for independent expert reviews of the agency's scientific products and programs ensures their credibility. Therefore, external scientific peer reviews have been and continue to be essential to strengthening scientific quality assurance for fishery conservation and management actions.

The Northeast Regional Stock Assessment Review Committee (SARC) meeting is a formal, multiple-day meeting of stock assessment experts who serve as a panel to peer-review tabled stock assessments and models. This review determines whether or not the scientific assessments are adequate to serve as a basis for developing fishery management advice.

The Red Hake Stock Structure Research Track SARC met in the Aquarium Conference Room at NOAA's Northeast Fisheries Science Center (NEFSC) in Woods Hole, MA during March 9 – 12th, 2020. The review committee was composed of three scientists appointed by the Center for Independent Experts (CIE): Manuel Hidalgo, Haritz Arrizabalaga, and Christophe Pampoulie, and was chaired by John Wiedenmann as a member of the New England Fisheries Management Council Scientific and Statistical Committee. The SARC was assisted by the NEFSC Stock Assessment Workshop (SAW) Chairman, James Weinberg, Michelle Traver, and Russ Brown. Documentation was prepared by the red hake Working Group (WG), and presentations were made primarily by the chair of the working group Dave Richardson (NEFSC), but other working group members presented material and contributed substantially to the discussions on various topics, including Steve Cadrin (U.Mass Dartmouth), Tim Miller (NEFSC), Rich McBride (NEFSC), Larry Alade (NEFSC), Toni Chute (NEFSC), and Kathy Sosebee (NEFSC). Alicia Miller, Jon Deroba, Brian Linton, and Charles Peretti from the NEFSC acted as rapporteurs throughout the meeting.

Review Panel and Review Activities

Dr. Christophe Pampoulie, Dr. Manuel Hidalgo and I served as CIE Reviewers, and Dr. John Wiedenmann chaired the meeting and drafted the summary report of the meeting, with input from all CIE reviewers. Our roles were to review the documents before the meeting, discuss with the authors their methods, findings, and interpretations, request any additional clarification, contribute to the summary report of the meeting, and report individually on our findings per the terms of reference. The Terms of Reference and my comments to each are given below.

General Comments

The Red Hake Stock Structure Working Group (WG) worked efficiently to conduct a very comprehensive study on a quite complex issue in a relatively short time period (considering that the first scoping webinar was conducted on the 4th of November 2019). The assigned Terms of Reference (ToRs) implied a diversity of tasks related to determination of stock structure and evaluation of stock status, as well as identification of research priorities. Thus, the WG incorporated a wide range of expertise to accommodate the multidisciplinary nature of the

tasks. In general, very useful and detailed background information was provided and the analyses were very well documented. I sincerely thank the WG chair and participants for the interesting presentations and fruitful discussions during the SARC meeting. I conclude that all Terms of Reference were successfully met, but the reference points developed in ToR 5 require additional sensitivity tests before they can be used for management advice.

I have a general comment on the geographical scope of the work conducted by the WG. Ideally, management units should reflect the real biological population structure of the resource (Reiss et al. 2009). This requires analyzing the available information at the scale of the whole species distribution (in fact, ToR 1 asked to review and summarize all relevant literature "in the northwest Atlantic"). However, the final aim of this work is to inform management units in U.S. waters, and several analyses used data mostly collected in U.S. waters, and are used to support specific stock structure hypotheses. It would have been ideal to have this type of information consistently throughout the whole distribution area, by including comparable data obtained in Canadian waters. The U.S. fall surveys do cover the Western Scotian Shelf, and those data were considered in the analyses. In addition, the WG has considered, to the extent possible, different sources of information from Canadian studies and surveys under ToR 2. They also looked at the Canadian Department of Fisheries and Oceans (DFO) Bay of Fundy survey data. This effort by the WG was very positive as it allows getting a broader perspective on red hake population structure.

The WG report did not include formal definitions of important terms such as "stock" or "population". I understand that, throughout the WG report, "stock" refers basically to "management unit", and "stock structure" refers to the delineation of such management units. In fact, they seem to have used the term "population structure" only when referring to the genetic population structure. Using the terms "stock" and "stock structure" is appropriate given the objective to determine whether the current management units within U.S. waters should remain, or should be modified (Waldman 2005). However, some of the arguments supporting the current delimitation of management units might be related to the biology of red hake, while others might be related to fisheries practice and/or management considerations. Ideally, because the management units should match the biological population structure, I think it is useful to make a clear separation between those arguments that inform about the biological population structure of the resource, from those others that do not. For this, it might have been practical to use the terms "stock structure" and "population structure" where appropriate. Although the WG did not do so, sometimes they did correctly indicate that some analyses where not informative about the biology of red hake, which was helpful.

Below are comments on each of the ToRs. I think that the WG correctly recommended maintaining the current two stock separation for red hake. However, this decision largely relies on the fact that there is no clear better alternative. Important uncertainties around red hake stock structure remain, which might have important implications on the success of the U.S. stock assessment and management process; thus further research is recommended in the future to try to improve the situation.

Specific Comments on each Term of Reference (ToR)

ToR1. Review and summarize all relevant literature on the existing stock structure of red hake in the northwest Atlantic.

This ToR was met satisfactorily.

Under this ToR, the WG provided a very useful chronology to understand the information that was available to inform stock designations over time.

Obviously, more information has become available through time, but it is also obvious that the population structure of red hake is not totally resolved yet and requires further research in the future. The fact that no tagging nor genetic information is available for red hake has been a strong handicap to properly evaluate its stock structure. Past decisions were partly based on biological information (e.g., meristics, distribution of the species, growth, etc.), but also other considerations (e.g., "similarity to silver hake, a species with substantial overlap in the small mesh fishery") that are not at all informative about red hake population structure, but are just practical considerations taken into account when deciding about management units.

ToR2. Identify and evaluate any new and/or existing data relevant to the stock structure of red hake including but not limited to the species' life history (i.e. spawning, distribution, abundance, growth, maturity and natural mortality), morphometrics, and genetics.

This ToR was met satisfactorily.

The WG did an excellent job in compiling all available information on red hake and conducted several analyses to shine light on the stock structure of red hake. Below are specific comments for each topic, although some comments are obviously related to other topics and ToRs.

Fishery Dependent Data

The WG successfully compiled and analyzed different sources of fishery dependent information, including landings, vessel trip reports, self-reported data, and observer programs. They also successfully interacted with fishermen to get better knowledge of the species and fishery dynamics.

While fishery dependent data can provide useful insights, they must also be interpreted with caution, as they are typically subject to potential sources of error and bias, which can be related to the sampling strategies followed, the level of sampling achieved, or changes in fishery dynamics due to regulations, multiple species of interest, observers onboard, etc. In this case, the observer coverage seems to have varied through time, reaching a relatively high coverage (up to 30% in the silver hake fishery) in the latest years. Some vessels provide detailed reports of their activity with haul-by-haul data. These seem to be reliable when compared to observer data, but are provided by a non-random sample of boats.

Maps of hauls containing red hake while targeting different species were used to illustrate that the distribution of red hake catch was patchy and discontinuous, with little to no catch around

the current stock boundary. The WG correctly noted that this spatial distribution of the catch should not be interpreted as the distribution of the species, since it is not a target species and it is caught in small mesh fisheries, which in some cases have restricted areas closed to fishing.

The WG found that most of the vessels had landed red hake from one or the other stock, but few vessels had landed red hake from both stocks. This might be a practical consideration to decide on management units, as it might simplify the monitoring of stock specific catches, but it clearly does not inform about the biological population structure of red hake.

On the other hand, the WG used Multivariate Auto-Regressive State-Space (MARSS) models to evaluate coherence in trends in catch and CPUE data across different fishing statistical areas. After considering several hypotheses of a single as well as multiple stock combinations, they found support for the two-stock model, with the two stock areas most often consistent with the current boundaries of the stock. This finding is interesting, but it was not discussed in much detail, which is partly justified because similar analyses were conducted based on fisheries independent survey data which should provide clearer insights on stock structure.

Trawl survey distributions

The WG presented analyses based on very comprehensive scientific trawl survey datasets. NEFSC trawl surveys started in 1963, with both spring and fall surveys occurring since 1968. Since then, the most important change relates to the change in boat and gear in 2009. This was accompanied by a calibration study (Miller et al. 2010), which allows one to use a longer time series in some of the analyses. Moreover, the WG had access to data from the Canadian Department of Fisheries and Oceans (DFO) trawl survey conducted in the Bay of Fundy to the eastern Scotian Shelf, which was useful for more complete insights.

Red hake spatial distribution maps based on the trawl surveys show a similar pattern to the one observed with the fishery data, with low density around the current two-stock boundary. This could be interpreted as supporting the two-stock hypothesis, but there is also room for alternative interpretations. In fact, the current stock boundary is a relatively shallow area which represents a poor habitat for red hake. Thus, a single population could spread throughout the species distribution range in proportion to the optimum habitat, and show a similar pattern to the one observed. It is certainly interesting to analyze the extent to which this poorer habitat around the current stock boundary may have acted as an oceanographic barrier to the mixing of red hake inhabiting northern and southern areas. West of Georges Bank there is a relatively deeper, more suitable area for red hake that might represent an ecological bridge (sensu Briscoe et al. 2017) between the north and the south. The analyses conducted by the WG suggest fastest spring habitat improvements might be occurring south and west of Georges Bank, so close to and across the stock boundary, which might be increasing connectivity between the two areas. It is also important to consider the dynamic nature of the whole U.S Northeast shelf ecosystem that is experiencing a rapid warming related to climate change, which might imply not only changes on red hake physical habitat preferences, but the organization of the whole ecosystem (Friedland et al. 2020).

Distribution maps showed a clear shift of red hake towards the northeast through time. Biomass decreased over time in southern areas and increased in northern areas. However, the mechanisms driving those observations remain unclear. Habitat model predictions show that the preferred habitat has also shifted in the same direction. Under a single population hypothesis, the observed shift in distribution could be due to a redistribution of the population according to changes in optimum habitat (“niche tracking”, sensu Erauskin-Extramiana et al.

2019). Alternatively, under a two-population hypothesis, it might be that the southern stock is becoming less productive while the northern one is becoming more productive, which could also be due to changes in habitat quality.

The application of the Management Unit Estimator (MUE, by Cope and Punt (2009)) was very interesting. This approach uses fine scale relative abundance data and applies a clustering method to group areas with similar population trends, under the assumption that areas that are clustered together behave as a single stock with the same population dynamics. I think this tool was very appropriate for the type of data that was available and the question at stake. The group appropriately tried to accommodate data from the Scotian Shelf in order to use data covering as much spatial distribution as possible, although this finally did not work so well as it required using data from shorter time periods.

According to the WG report, the application of the MUE provided meaningful support for the existing stock structure. Moreover, when the MUE was a priori set to 2 stocks, the current stock definitions are the best, with meaningful-to-notable support. While I agree with both interpretations, I have some reservations on the first one, since I think that the results do provide support for the two-stock hypothesis compared to the three or four stocks hypotheses, but no formal comparison is made between the two-stock and the single stock hypotheses. The clustering method in Cope and Punt (2009) considers k=2 or more stocks, so it does not allow for a formal comparison with the single stock hypothesis. I understand that if the MUE provides some support for the two-stock hypothesis, it implies that some structure exists in the dataset, and this can be interpreted as some indirect support for the two-stock hypothesis against the single stock hypothesis, but yet it is not a formal statistical comparison.

Age and Growth

The working group conducted substantial analyses on size at age datasets from the scientific trawl surveys. Data were collected during two different time periods (1970-1985 and 2008-2018, respectively). The most remarkable result with respect to stock structure was that red hake seems to grow more in northern areas compared to southern areas, which could support the two-stock hypothesis. The difference was more pronounced in the early period, but apparently red hake mis-identification in that period could be affecting this result, at least to some extent. However, the geographical differences seem clear as the misidentification does not affect the most recent period. The WG did not perform rigorous statistical tests on geographical differences between growth curves fit to northern and southern length at age datasets, but the data were shown in different ways and the patterns were quite clear. Moreover, the Management Unit Estimator was also applied to length at age data in different regions, providing support for the two-stock hypothesis. The comment I made on the MUE in the previous section is also applicable here, as the approach does not allow one to formally compare the single stock and the two-stock hypotheses.

In addition, although the analyses performed are informative for the objectives of the WG, I think that in this case there is enough data (even considering the most recent period, with around 10,000 individuals aged) to perform a more comprehensive statistical analysis and try to understand what is structuring the differences in growth. Differences in growth can be due to genetic differences, or phenotypic responses to variation in local environmental factors such as temperature and food availability (Weatherley and Gill 1987). Spatial variation in growth has been observed in both demersal and pelagic species (Williams et al. 2012). In the case of red hake, in order to interpret growth in relation to stock structure, it would be useful to see

whether growth shows some kind of spatial gradient (e.g., in latitude), rather than just differing between the north and the south.

Another important finding was that a length truncation occurred through time in all areas. This is a very clear and interesting pattern that might be indicative of different processes. The most likely hypothesis is that growth patterns are changing due to a warming scenario, with smaller asymptotic sizes in recent years. Given the exploitation history of the species (red hake has not been a main target species for many decades and catches remained far below historical high values), it is unlikely that fisheries induced the observed age truncation. However, the WG observed a very prominent increase in red hake predation over time, which could be playing a role.

Otolith microchemistry

There was limited information available on otolith chemistry as only a pilot project was conducted with a limited number of samples from a single year (2011). The study area presents strong oceanographic gradients and the selected markers were adequate as natural environmental tags. No differences in the chemistry of the otolith nucleus in individuals from Gulf of Maine and Mid Atlantic Bight suggests a common source, supporting the single population hypothesis. However, it is strange not to find differences in the edge of the otoliths as fish were collected in substantially different environments. This could imply a recent migration to those areas, but mostly calls for caution on the interpretation of the nucleus data until the study is expanded. I think that otolith chemistry can be a very promising research avenue in the future, and it is worth expanding the sample size to several years. It is also worth exploring other markers such as oxygen isotopes (LeGrande and Schmidt 2006). The WG made a correct interpretation of the otolith chemistry study.

Spawning, larvae and Young-of-the-Year (YOY)

Under these sections the working group presented a large amount of very useful information. They did a good job analyzing the available data from different sources and linking everything together to inform the potential stock structure of red hake. In summary, they provided information on where the larvae are found, what are the potential sources of these larvae and where the YOY are found.

Data from two very comprehensive and consistent larval sampling programs (MARMAP and EcoMon surveys) where used, with observations since 1977. However, reliable red hake larval identification protocols became available only recently, and a large effort was conducted to reidentify red hake larvae using this new protocol but it was not possible to process all the historical information, so a set of 6 years distributed between 1985 and 2013 were processed. The species re-identification effort was quite substantial and the criteria to select years along the time series were adequate, allowing for very useful insights into population structure.

Most of the larvae are observed in Georges Bank and Southern New England, primarily in the southern stock area, but some larvae are also found in northern stock areas, including the Gulf of Maine. Limited shifts in larval distribution over time were observed, contrasting with the distribution shifts observed for larger red hake in trawl surveys.

Larval transport and drifter models suggest that most spawning occurs on Georges Bank and Southern New England, but also some larvae are locally produced in the Gulf of Maine. Most of the small YOY (1-10 cm) were also largely found around Georges Bank in fall, especially in the northern and western parts. Interestingly, slightly larger YOY (11-18cm) are mostly distributed

in deep areas of the Gulf of Maine, north of Georges Bank, which suggests a potential migration of YOY born in Georges Bank to overwintering areas in the Gulf of Maine.

The working group presented a potential conceptual model where adults overwintering in the Gulf of Maine would perform a spawning migration to Georges Bank spawning areas. Larvae would be drifted/transported to the shallow Georges Bank and then the YOY would perform the overwintering migration to the Gulf of Maine. This conceptual model could support the single stock hypothesis, but could also support multiple stocks with a single spawning area. Unfortunately, there is no evidence of site fidelity to northern and southern feeding areas for individuals sharing spawning grounds in the Georges Bank, which could help discriminate between the two hypotheses. In the absence of tagging experiments, additional genetic and/or otolith chemistry analyses might provide more insight on this.

One additional piece of the puzzle deserves further attention. In recent years, YOY abundance has largely increased in the northern Gulf of Maine. Data from the DFO summer survey in the Bay of Fundy also show a similar increase in YOY during the last decade. Larval connectivity models suggest that juveniles that recruit to the Gulf of Maine are likely spawned in the Gulf of Maine, although there is some possibility that they come from the Bay of Fundy/Western Scotian Shelf (if there was spawning there), and are unlikely to come from the Georges Bank. Unfortunately, there is no clear information about larval distribution in the Bay of Fundy to assess on potential spawning in that area that could explain at least part of the recruitment in the Gulf of Maine. In addition, the EcoMon larval sampling in August has not surveyed the northern Gulf of Maine since 2013 due to limited ship time availability. I recommend addressing this information gap in the near future and explore possible spawning in that region to see if patterns coincide with those observed in YOY.

Application of An Index Method

The working group used the AIM model as a potential way to provide insights on red hake stock structure. If the appropriate stock structure is chosen and if exploitation rates are driving population trends, then the AIM model should be significant. The working group tested multiple alternative stock structures, including one, two and three stock hypotheses. However, none of the models were significant, thus this approach did not provide useful insights into stock structure. Considering that a wide range of possible stock structures where considered, non-significance of the AIM models might suggest that fishing mortality is not driving abundance of red hake. In fact, red hake consumption (by predators) was estimated to be six times higher than catch levels during the period 2000-2010 (NEFSC 2011), and the consumption estimates (limited to a suite of predators caught during the trawl surveys) are probably conservative. Alternatively, non-significance of the AIM models might also imply that red hake stock structure is different from the range of possibilities tested.

ToR3. Recommend the most likely biological stock structure among a set of alternatives from TOR2. Consider the current management unit as null hypothesis.

This ToR was met satisfactorily.

The WG followed an interdisciplinary approach to stock identification, and they considered the available information insufficient to reject the null hypothesis of two stocks, a northern stock

that encompasses the northern Georges Bank and the Gulf of Maine, and a southern stock that encompasses the southern New England and the southern Georges Bank. This decision was supported by the observed phenotypic differences (in growth, meristics and otolith morphology), the consistency of relative abundance indices in those areas (as determined by the MUE), and fisheries dynamics suggesting easy separation of basic fishery statistics among the two stocks. Obviously, the latter does not provide biological support, but it is a practical consideration. The WG also did a good job in confronting the information supporting the current stock structure with the one that would support the alternative single stock hypothesis. This mostly includes the information on the early life history and spawning, but also the microchemistry study, in spite of its limitations. The WG also suggested a reasonable potential life history that would fit under the single stock hypothesis, but there was no real evidence to support this and reject the null hypothesis.

The WG provided a summary table with their findings supporting alternative stock structure hypotheses (one, two or more stocks). Life history data, and specifically growth differences, had a relatively strong impact on the final decision. However, life history data alone are typically not enough for properly characterizing stock structure as size at age is not necessarily linked to genetic variation but can vary across geographic scales due to a variety of other reasons including environmental influence (McBride 2014). In this situation, the consistency of the relative abundance indices in the two areas was critical to support the two-stock hypothesis.

The WG decision was obviously related to the specific way the ToR was drafted. For instance, if instead of having to decide whether there is sufficient information to reject the null hypothesis, the request was to propose the best stock structure given the available information (without a null hypothesis), the single stock hypothesis could have been weighted more.

The potential migration pattern they suggested under the single stock hypothesis implied a spawning migration from the Gulf of Maine to spawning areas around the Georges Bank, and a juvenile migration from Georges Bank back to the Gulf of Maine. This is a way to reconcile the finding of a single major spawning area for both stocks, and it is consistent with some observations within U.S. waters (e.g., the distribution of small and larger YOY over time). However, I think that the group could have further entertained alternative hypotheses such as potential links between the northern stock to spawning areas in Canadian waters, given the increase in YOY observed during recent years in both the northern Gulf of Maine and the Bay of Fundy/Western Scotian Shelf. This could have shed light on the potential need for further joint transboundary research and eventually assessments.

The lack of tagging and genetic data complicated the interpretation of the available information and could not support the potential spawning migration hypothesis. A better understanding of the homing behavior to feeding areas after spawning would also help shape red hake population structure and dynamics: do offspring of northern red hake recruit mostly in northern areas after spawning in Georges Bank, or not?

All in all, the work conducted by the WG highlighted the complexity of the issue and the need for further research to resolve it. In this situation, while I fully agree with the conclusion by the WG to go ahead with the current stock structure, I recommend conducting some simple explorations on the effects of alternative population structures on the perception of red hake stock status (Arrizabalaga et al. 2007). Basically, this could be explored in a rather simple way by considering that all observations across the whole spatial distribution belong to the same management unit (single stock hypothesis). Assessing stock status under this alternative

scenario would provide a rough idea of the importance to resolve stock structure, which could help prioritize future research in the short and longer terms (ToR 6).

ToR4. Evaluate existing experimental data on survey catchability of red hake. Examine the sufficiency of catchability data and, if appropriate, incorporate the catchability estimates into the assessment.

This ToR was met satisfactorily.

The analyses presented under this ToR where very comprehensive and well designed. They were based on very substantial and novel datasets, followed by very rigorous state of the art statistical analyses that allowed the WG to make robust conclusions.

Since 2009, the gear used in the bottom trawl surveys conducted by the Northeast Fisheries Science Center is equipped with a rockhopper sweep, which can be consistently applied over a wide range of bottom types, but allows for some fish to pass underneath the net. In order to estimate catchability of the gear used for the surveys, a paired gear study was conducted where two gears, one with the rockhopper sweep and the other with a chain sweep were towed side by side. The gear with the chain sweep was assumed to have a 100% catchability, and the relative efficiency of the two gears was used to estimate the catchability of the survey gear. This approach has been peer reviewed and used to estimate catchability and swept area biomass for flatfish in the Northwest Atlantic (Miller 2013). Red hake is the first roundfish that went through the same process, and the catchability estimates can now be used in the assessment.

The WG also used images from the HABCAM survey to estimate red hake abundance and compare it with trawl survey estimates, which can also inform about the trawl survey gear catchability under the assumption that red hake does not avoid the gear used to tow the habitat camera. A single year (2015) of data was used in this case, and the WG reported some problems in the HABCAM dataset as red hake and spotted hake could not always be properly identified in the images. Anyway, the WG noted that when considering ratios of the two species consistent with the ones observed in the trawl surveys, the catchability estimates were comparable to those of the chain sweep study. This is comforting and supports the use of the chain sweep study results in the assessment, but as the WG correctly concluded, these estimates should not be used as the primary source of catchability information.

The WG also presented information from a net wing spread study aimed at evaluating the impact of wing spread on catchability. The study was designed to test the hypothesis that gear efficiency declines when the net is overspread or underspread, which typically happens in deep and shallow tows, respectively. The results presented by the WG suggest that this is not the case for red hake. However, they also noted that this is the first species analyzed and a multispecies evaluation of the data is ongoing, which might provide additional light on the effects of wingspread on catchability. Anyway, the WG did recommend that future calculations of swept area biomass should explore directly accounting for measured wingspread rather than applying the average wingspread to all tows, which I think is a reasonable recommendation.

ToR5. Apply the existing assessment model framework to the stock structure based on TOR 3 and 4 to ensure its utility in subsequent management track assessments. Evaluate existing reference points. Consider alternate assessment approaches if existing model framework does not perform well, and consider alternate reference points as needed.

This ToR was met satisfactorily, although the estimated alternate reference points require additional sensitivity analyses before they can be used for management.

The WG applied the existing AIM model to the northern and southern stocks. The AIM model tests whether fishing mortality drives the trend in biomass and calculates a relative fishing mortality that allows the stock to replace itself (replacement ratio of 1), which is considered a proxy for F_{MSY}. A randomization test is used to evaluate the significance of the models. For both stocks, the AIM models were not significant and provided very low F_{MSY} estimates, thus the WG determined that this approach should not be used in the future stock assessments. I agree with this decision, and also think that the lack of significance of the models, in itself, is an interesting result as it indicates that, under the assumed stock structure, fishing mortality is not driving abundance of the stocks. If this is true, some management decisions could have little impact on the stocks.

Following the ToR, the WG considered an alternative method for calculating reference points for red hake. They proposed spawning potential ratio (SPR) reference points, which are widely used as proxies for MSY-based reference points. The WG proposed $F_{40\%}$ as a proxy for F_{MSY}. $F_{40\%}$ is the fishing mortality rate that would reduce the spawning potential to 40% with respect to the spawning potential in the absence of fishing. The associated spawning biomass per recruit can be multiplied by the number of recruits to obtain $SSB_{40\%}$, the proxy for SSB_{MSY}. SPR based reference points are technically sound and the 40% proxy is reasonable and commonly used in a wide range of stocks (Clark 1991, Miller and Rago 2012). Thus, I think that the general approach can be useful in future stock assessment and management applications. However, there were some concerns around the estimates, including that some SSB to $SSB_{40\%}$ ratios were unrealistically high. Thus, the particular $F_{40\%}$ and $SSB_{40\%}$ estimates for north and south red hake should not be directly used to set catch advice in the short term, until further sensitivity analyses are conducted with respect to some of the assumptions, such as natural mortality, selectivity and recruitment.

The WG assumed an M value of 0.4 yr^{-1}, but M might have changed over time and variations in M can affect the reference point estimates. Legault and Palmer (2016) suggest considering the amount of empirical evidence around time varying M before considering a range of alternatives in the assessment. Friedland et al. (2020) suggest that the whole U.S. Northeast Shelf ecosystem has reorganized and the niche overlap between species might have increased. On the other hand, red hake consumption estimates suggest that there might have been a change in natural mortality over time (NEFSC 2011). These consumption estimates probably represent minimum consumption rates and are subject to considerable uncertainty (i.e., because only a part of the predators of red hake are sampled). Thus, it is hard to confirm whether time trends are really significant, but the data suggests this is a real possibility and thus I think it is worth investigating the effects of alternative M values in the reference points.

Another important assumption is that the fishery has a knife edge selectivity pattern, essentially assuming that fish of ages greater than 1 are equally selected. Variations in the selectivity pattern typically impact SPR reference points, so conducting a sensitivity analysis around alternative selectivity patterns is recommended. If the real selectivity pattern does not fully select all those ages, reference points are expected to increase. On the other hand, the WG noted that some YOY associate with scallops for shelter. If such YOY are bycaught in significant quantities in the scallop or other fisheries, the selectivity at age 0 might be higher than the one assumed by the WG and additional sensitivity analyses should also be conducted around these alternative plausible selectivities.

Finally, the WG used the average recruitment value for the most recent period (2009-2019) to estimate $SSB_{40\%}$. This is justified because the survey is very consistent during these years, but it also represents a short time series, and it is unclear whether the average recruitment during this period reflects the productivity of the stock at equilibrium or not. A longer time series of recruitment would allow investigating whether alternative recruitment values are worth considering, which can significantly impact $SSB_{40\%}$ estimates and the stock status determination.

In spite of the concerns around the reference points, exploitation rates for both stocks are very low and overfishing is not likely occurring. Survey data also indicate that the biomass in northern areas is increasing, thus it is also likely that the northern stock is not overfished. Conversely, the dynamics in the southern stock are quite different and the status of the stock is more uncertain.

Survey data indicate that red hake in southern areas has decreased very substantially and remains at very low values since the 1990s. The volume of catches supported by the stock during the last three decades is also minor compared to historical values observed before the late 1970s when distant water fleets left the area. These large catches, peaking at around 100,000 MT in 1966, probably influenced the decrease in biomass. However, the stock has not recovered to historical biomass levels after the drastic reduction in catches during several decades. This suggests that a decrease in stock productivity might have occurred, which could be due to a decrease in habitat quality, an increase of predation by other species leading to a higher natural mortality, or other reasons. In any case, it would be interesting to test the sensitivity of $SSB_{40\%}$ to higher recruitment values reflecting potentially higher productivity.

The lack of stock recovery despite very low exploitation also suggests that exploitation is not driving biomass trends. In fact, consumption of red hake to catch of red hake ratios (NEFSC 2011) indicate that the effect of fishing mortality might be negligible compared to the effect of M, and this would be consistent with the lack of convergence of the AIM model. Under these circumstances, it might not be possible to recover historical productivity levels by further reducing exploitation on southern red hake, and it might not be possible to sustain historical catch levels under the current productivity regime. Thus, although I recommend performing sensitivity analysis of $SSB_{40\%}$ with respect to higher equilibrium recruitment levels, it might not be reasonable to consider these reference points as a management target. In fact, the risk to further impair recruitment should be strongly considered before any catch increase is recommended.

In the future, additional methods currently used for setting catch advice for data poor stocks could be explored (e.g., Jardim et al. 2015). Additional understanding on the stock dynamics and insights on management options might also be obtained by moving beyond single species assessment models, although moving towards an ecosystem approach to fisheries management is not straightforward (Mace 2001).

ToR6. Identify gaps in the existing research with respect to red hake stock structure. Develop a prioritized list of research recommendations to address these gaps. Comment on the feasibility and time horizon of the proposed research recommendations.

This ToR was met satisfactorily.

The WG identified main knowledge gaps and developed a prioritized list of recommendations to address them, including approximate cost estimates and comments on feasibility. There are two sets of recommendations, one with red hake specific recommendations and another one with general recommendations.

I agree with the general set of recommendations (all considered high priority), and especially on the need to have full spatial coverage during the ichthyoplankton surveys as they are of utmost importance to identify potential spawning areas. Failure to fully cover the northernmost strata in U.S. waters contributed to the uncertainty on red hake stock structure, and it is important to correct this in the future, as well as to avoid a similar situation for other species when their stock structure is discussed.

As for red hake specific recommendations, a modern genetic study is a real priority as it can really shed light into the biological population structure of red hake. The NEFSC scientists have great access to reference samples (individuals of known origin such as larvae or spawning adults) throughout the distribution range which can enable understanding of how populations are structured, and build genetic tools to assign individuals to origin. These tools can be applied to a broad set of other non-reference samples (mixing aggregations) to understand how the different populations eventually mix in space and time. The proposed sample size (n=500) might be a bit short if the objective is to check for temporal stability of the observed patterns, as it requires sampling from different seasons and years. One option could be to prioritize complete sampling coverage in a single year to have a good "snapshot" of population structure, and then consider further sampling in following years. In case of budget limitations, they could also first focus on analyzing the most distant samples, where the chance of observing genetic differences is highest.

However, genetic studies do not always provide all the answers needed for stock assessment and management (Waples et al. 2008). For instance, small amounts of genetic mixing might prevent genetic differentiation, but different contingents might still deserve differential management. Thus, I would recommend conducting the genetic analysis coupled with the otolith chemistry study and the analysis of other natural tags, as these markers provide complementary information on different aspects of the biology (e.g., ancestry vs environment). Moreover, additional insight arises when different analyses are conducted on the same individuals (although obviously this is not possible on larvae). But even if they are not conducted on the same individuals, conducting parallel studies over the same time period and spatial distribution can be useful (Waldman 1999, Cadrin et al. 2013).

The WG considered the otolith chemistry study a medium priority, but I think it is a high priority. There is environmental contrast throughout the study area for potentially successful application of this methodology. The study conducted in the past was preliminary in nature and it is worth investing further efforts on this topic. Moreover, if the lack of difference in red hake otolith

nucleus signature from northern and southern areas is confirmed, it could support the single spawning area hypothesis.

References

Arrizabalaga, H., López-Rodas, V., Costas, E. and González-Garcés, A. (2007) Use of genetic data to assess the uncertainty in stock assessments due to the assumed stock structure: The case of albacore (Thunnus alalunga) from the Atlantic Ocean. *Fishery Bulletin* **105**: 140-146.

Briscoe, D.K., Hobday, A.J., Carlisle, A., Scales, K., Eveson, J.P., Arrizabalaga, H., Druon, J.N. and Fromentin, J.-M. (2017) Ecological bridges and barriers in pelagic ecosystems. *Deep Sea Research Part II: Topical Studies in Oceanography* **140**: 182-192.

Cadrin, S.X., Kerr, L.A. and Mariani, S. (eds) (2013) Stock identification methods: applications in fishery science. Academic Press, 589 pp.

Clark, W.G. (1991) Groundfish exploitation rates based on life history parameters. *Canadian Journal of Fisheries and Aquatic Sciences* **48**: 734-750.

Cope, J.M. and Punt, A.E. (2009) Drawing the lines: resolving fishery management units with simple fisheries data. *Canadian Journal of Fisheries and Aquatic Sciences* **66**: 1256-1273.

Erauskin-Extramiana, M., Arrizabalaga, H., Cabré, A., Coelho, R., Rosa, D., Ibaibarriaga, L. and Chust, G. (2019) Are shifts in species distribution triggered by climate change? A swordfish case study. *Deep Sea Research Part II: Topical Studies in Oceanography*: 104666.

Friedland, K.D., Langan, J.A., Large, S.I., Selden, R.L., Link, J.S., Watson, R.A. and Collie, J.S. (2020) Changes in higher trophic level productivity, diversity and niche space in a rapidly warming continental shelf ecosystem. *Science of The Total Environment* **704**: 135270.

Jardim, E., Azevedo, M. and Brites, N.M. (2015) Harvest control rules for data limited stocks using length-based reference points and survey biomass indices. *Fisheries Research* **171**: 12-19.

Legault, C.M. and Palmer, M.C. (2016) In what direction should the fishing mortality target change when natural mortality increases within an assessment? *Canadian Journal of Fisheries and Aquatic Sciences* **73**: 349-357.

LeGrande, A.N. and Schmidt, G.A. (2006) Global gridded data set of the oxygen isotopic composition in seawater. *Geophysical research letters* **33.**

Mace, P.M. (2001) A new role for MSY in single-species and ecosystem approaches to fisheries stock assessment and management. *Fish and Fisheries* **2**: 2-32.

McBride, R.S. (2014) The continuing role of life history parameters to identify stock structure In: *Stock Identification Methods*.S.X. Cadrin, L.A. Kerr and S. Mariani (eds) Elsevier, pp. 77-107.

Miller, T. and Rago, P. (2012) Empirical exploration of feasible bounds on butterfish stock size and fishing mortality rates, 1975-2011. *Report to the Mid-Atlantic Fishery Management Council Scientific and Statistical Committee, May 2, 2012.*

Miller, T.J. (2013) A comparison of hierarchical models for relative catch efficiency based on paired-gear data for U.S. Northwest Atlantic fish stocks. *Canadian Journal of Fisheries and Aquatic Sciences* **70**: 1306-1316.

Miller, T.J., Das, C., Politis, P.J., Miller, A.S., Lucey, S.M., Legault, C.M., Brown, R.W. and Rago, P.J. (2010) Estimation of Albatross IV to Henry B. Bigelow calibration factors. Northeast Fisheries Science Center Ref. Doc. 10-05. 230 pgs.

Northeast Fisheries Science Center. 2011. 51st Northeast Regional Stock Assessment Workshop (51st SAW) Assessment Report. U.S. Dept Commer, Northeast Fish Sci Cent Ref Doc. 11-02; 856 p.

Reiss, H., Hoarau, G., Dickey-Collas, M. and Wolff, W.J. (2009) Genetic population structure of marine fish: Mismatch between biological and fisheries management units. *Fish and Fisheries* **10**: 361-395.

Waldman, J.R. (1999) The importance of comparative studies in stock analysis. *Fish. Res.* **43**: 237-246.

Waldman, J.R. (2005) Definition of stocks: an evolving concept In: *Stock identification methods*.S.X. Cadrin, K. Friedland and J. Waldman (eds) Elsevier, pp. 7-16.

Waples, R.S., Punt, A.E. and Cope, J.M. (2008) Integrating genetic data into management of marine resources: how can we do it better? *Fish and Fisheries* **9**: 423-449.

Weatherley, A.H., Gill, H. and Casselman, J.M. (1987) Biology of fish growthAcademic press, Pages pp.

Williams, A.J., Farley, J.H., Hoyle, S.D., Davies, C.R. and Nicol, S.J. (2012) Spatial and sex-specific variation in growth of albacore tuna (Thunnus alalunga) across the South Pacific Ocean. *PLoS ONE* **7.**

Appendix I: Materials provided for review

- Northeast Fisheries Science Center. 2011. 51st Northeast Regional Stock Assessment Workshop (51st SAW) Assessment Report. U.S. Dept. Commer., Northeast Fish Sci Cent Ref Doc. 11-02; 856 p.
- DRAFT REPORT. Red Hake Stock Structure Working Group. By the Northeast Fisheries Science Center, 11 February 2019.
- Appendix 1. Informing spatial structure of red hake (*Urophycis chuss*) stocks and the fleets that fish for them. Authors: Andrew Jones, Anna Mercer, David Richardson.
- Appendix 2. Application of the management unit estimator to red hake trawl survey data. David Richardson.
- Appendix 3. What size at age says about red hake stock structure. Richard S. McBride, Woods Hole Laboratory, NOAA Fisheries. DRAFT, not for wide distribution, 2 January 2020.
- Appendix 4. An empirical approach to assessing northern and southern red hake. Timothy J. Miller.
- Figures_RedHakeSSWG_11February2020.docx
- Tables_RedHakeSSWG_11February2020.docx
- Application of An Index Method (AIM) to Data Rich Situations: Can Simple Methods Capture Major Features of Complex Assessments? Paul J. Rago and Christopher M. Legault.
- Red Hake Stock Structure Research Track Terms of Reference (v. 2/27/2020).
- SARC 54 PANEL SUMMARY REPORT. 54th Northeast Regional Stock Assessment Workshop (SAW 54) Stock Assessment Review Committee (SARC) Meeting 5 - 9 June 2012 Northeast Fisheries Science Center Woods Hole, Mass.
- Text Final Report_Red Hake. SSWG. 11February2020.docx
- Northeast Fisheries Science Center Reference Document 11-01. 51st Northeast Regional Stock Assessment Workshop (51st SAW): Assessment Summary Report (2nd Edition). Aug. 2011.
- Northeast Fisheries Science Center Reference Document 12-18. 54th Northeast Regional Stock Assessment Workshop (54th SAW) Assessment Report. Dec. 2012.
- Northeast Fisheries Science Center Reference Document 18-02. 2017 Northern and Southern Silver Hake and Red Hake Stock Assessment Update Report. by Larry Alade and Michele Traver.
- Red Hake Stock Structure Research Track Assessment Peer Review Meeting. Clark Conference Room, NEFSC, Woods Hole, MA. March 9-12, 2020. Meeting Agenda.
- Various ppt Powerpoint presentations, covering each WG TOR for this meeting.

Appendix 2: Performance work statement

Performance Work Statement (PWS)
National Oceanic and Atmospheric Administration (NOAA)
National Marine Fisheries Service (NMFS)
Center for Independent Experts (CIE) Program
External Independent Peer Review

Red Hake Stock Structure Research Track

Background

The National Marine Fisheries Service (NMFS) is mandated by the Magnuson-Stevens Fishery Conservation and Management Act, Endangered Species Act, and Marine Mammal Protection Act to conserve, protect, and manage our nation's marine living resources based upon the best scientific information available (BSIA). NMFS science products, including scientific advice, are often controversial and may require timely scientific peer reviews that are strictly independent of all outside influences. A formal external process for independent expert reviews of the agency's scientific products and programs ensures their credibility. Therefore, external scientific peer reviews have been and continue to be essential to strengthening scientific quality assurance for fishery conservation and management actions.

Scientific peer review is defined as the organized review process where one or more qualified experts review scientific information to ensure quality and credibility. These expert(s) must conduct their peer review impartially, objectively, and without conflicts of interest. Each reviewer must also be independent from the development of the science, without influence from any position that the agency or constituent groups may have. Furthermore, the Office of Management and Budget (OMB), authorized by the Information Quality Act, requires all federal agencies to conduct peer reviews of highly influential and controversial science before dissemination, and that peer reviewers must be deemed qualified based on the OMB Peer Review Bulletin standards[1]. Further information on the Center for Independent Experts (CIE) program may be obtained from www.ciereviews.org.

Scope

The Northeast Regional Stock Assessment Review Committee (SARC) meeting is a formal, multiple-day meeting of stock assessment experts who serve as a panel to peer-review tabled stock assessments and models. The SARC peer review is the cornerstone of the Northeast Stock Assessment Workshop (SAW) process, which includes assessment development, and report preparation (which is done by SAW Working Groups or Atlantic States Marine Fisheries Commission (ASMFC) technical committees), assessment peer review (by the SARC), public presentations, and document publication. This review determines whether or not the scientific assessments are adequate to serve as a basis for developing fishery management advice. Results provide the scientific basis for fisheries within the jurisdiction of NOAA's Greater Atlantic Regional Fisheries Office (GARFO).

1 http://www.cio.noaa.gov/services_programs/pdfs/OMB_Peer_Review_Bulletin_m05-03.pdf

The purpose of this meeting will be to provide an external peer review of red hake stock structure. The requirements for the peer review follow. This Performance Work Statement (PWS) also includes: **Appendix 1**: TORs for the research track, which are the responsibility of the analysts; **Appendix 2:** a draft meeting agenda; **Appendix 3:** Individual Independent Review Report Requirements; and **Appendix 4:** Peer Reviewer Summary Report Requirements.

Requirements
NMFS requires three reviewers under this contract (i.e. subject to CIE standards for reviewers) to participate in the panel review. The chair, who is in addition to the three reviewers, will be provided by either the New England or Mid-Atlantic Fishery Management Council's Science and Statistical Committee; although the chair will be participating in this review, the chair's participation (i.e. labor and travel) is not covered by this contract.

Each reviewer will write an individual review report in accordance with the PWS, OMB Guidelines, and the TORs below. All TORs must be addressed in each reviewer's report. No more than one of the reviewers selected for this review is permitted to have served on a SARC panel that reviewed this same species in the past. The reviewers shall have working knowledge and recent experience in the evaluation of biological and ecological data commonly used in stock delineation for marine fishes including but not limited to life history traits, morphometric data, seasonal and spawning distribution data, otolith microchemistry data, and genetics. In addition, knowledge and experience with data limited assessment and population dynamics would be valuable.

Tasks for Reviewers

- Review the background materials and reports prior to the review meeting
    - Two weeks before the peer review, the Assessment Process Lead will electronically disseminate all necessary background information and reports to the CIE reviewers for the peer review.
- Attend and participate in the panel review meeting
    - The meeting will consist of presentations by NOAA and other scientists, stock assessment authors and others to facilitate the review, to provide any additional information required by the reviewers, and to answer any questions from reviewers
- Reviewers shall conduct an independent peer review in accordance with the requirements specified in this PWS and TORs, in adherence with the required formatting and content guidelines; reviewers are not required to reach a consensus.
- Each reviewer shall assist the SARC Chair with contributions to the Peer Reviewer Summary Report
- Deliver individual Independent Reviewer Reports to the Government according to the specified milestone dates

- This report should explain whether each research track Term of Reference was or was not completed successfully during the SARC meeting, using the criteria specified below in the "Tasks for SARC panel."
- If any existing Biological Reference Points (BRP) or their proxies are considered inappropriate, the Independent Report should include recommendations and justification for suitable alternatives. If such alternatives cannot be identified, then the report should indicate that the existing BRPs are the best available at this time.
- During the meeting, additional questions that were not in the Terms of Reference but that are directly related to the assessments and research topics may be raised. Comments on these questions should be included in a separate section at the end of the Independent Report produced by each reviewer.
- The Independent Report can also be used to provide greater detail than the Peer Reviewer Summary Report on specific stock assessment Terms of Reference or on additional questions raised during the meeting.

Tasks for Review panel

- During the SARC meeting, the panel is to determine whether each research track Term of Reference (TOR) was or was not completed successfully. To make this determination, panelists should consider whether the work provides a scientifically credible basis for developing fishery management advice. Criteria to consider include: whether the data were adequate and used properly, the analyses and models were carried out correctly, and the conclusions are correct/reasonable. If alternative assessment models and model assumptions are presented, evaluate their strengths and weaknesses and then recommend which, if any, scientific approach should be adopted. Where possible, the SARC chair shall identify or facilitate agreement among the reviewers for each research track TOR.
- If the panel rejects any of the current BRP or BRP proxies (for B_{MSY} and F_{MSY} and MSY), the panel should explain why those particular BRPs or proxies are not suitable, and the panel should recommend suitable alternatives. If such alternatives cannot be identified, then the panel should indicate that the existing BRPs or BRP proxies are the best available at this time.
- Each reviewer shall complete the tasks in accordance with the PWS and Schedule of Milestones and Deliverables below.

Tasks for SARC chair and reviewers combined:

Review the Report of the Red Hake Stock Structure Working Group.

The SARC Chair, with the assistance from the reviewers, will write the Peer Reviewer Summary Report. Each reviewer and the chair will discuss whether they hold similar views on each research track Term of Reference and whether their opinions can be summarized into a single conclusion for all or only for some of the Terms of Reference of the SAW. For terms where a similar view can be reached, the Peer Reviewer Summary Report will contain a summary of such opinions. In cases where multiple and/or differing views exist on a given Term of Reference, the Peer Reviewer Summary

Report will note that there is no agreement and will specify - in a summary manner – what the different opinions are and the reason(s) for the difference in opinions.

The chair's objective during this Peer Reviewer Summary Report development process will be to identify or facilitate the finding of an agreement rather than forcing the panel to reach an agreement. The chair will take the lead in editing and completing this report. The chair may express the chair's opinion on each research track Term of Reference, either as part of the group opinion, or as a separate minority opinion. The Peer Reviewer Summary Report will not be submitted, reviewed, or approved by the Contractor.

If any existing Biological Reference Points (BRP) or BRP proxies are considered inappropriate, the Peer Reviewer Summary Report should include recommendations and justification for suitable alternatives. If such alternatives cannot be identified, then the report should indicate that the existing BRP proxies are the best available at this time.

Foreign National Security Clearance
When 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 reviewers who are non-US citizens. For this reason, the reviewers shall provide requested information (e.g., first and last name, contact information, gender, birth date, country of birth, country of citizenship, country of permanent residence, country of current residence, dual citizenship (yes, no), passport number, country of passport, travel dates.) to the NEFSC Assessment Process Lead 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/ and http://deemedexports.noaa.gov/compliance_access_control_procedures/noaa-foreign-national-registration-system.html. The contractor is required to use all appropriate methods to safeguard Personally Identifiable Information (PII).

Place of Performance
The place of performance shall be at the contractor's facilities, and at the Northeast Fisheries Science Center in Woods Hole, Massachusetts.

Period of Performance
The period of performance shall be from the time of award through April 30, 2020. Each reviewer's duties shall not exceed **14** days to complete all required tasks.

Schedule of Milestones and Deliverables: The contractor shall complete the tasks and deliverables in accordance with the following schedule.

Schedule	Milestones and Deliverables
Within 2 weeks of award	Contractor selects and confirms reviewers
Approximately 2 weeks later	Contractor provides the pre-review documents to the reviewers
March 9-12, 2019	Panel review meeting
Approximately 2 weeks later	Contractor receives draft reports
Within 2 weeks of receiving draft reports	Contractor submits final reports to the Government

* The Peer Reviewer Summary Report will not be submitted to, reviewed, or approved by the Contractor.

Applicable Performance Standards
The acceptance of the contract deliverables shall be based on three performance standards:
(1) The reports shall be completed in accordance with the required formatting and content (2) The reports shall address each TOR as specified (3) The reports shall be delivered as specified in the schedule of milestones and deliverables.

Travel
All travel expenses shall be reimbursable in accordance with Federal Travel Regulations (http://www.gsa.gov/portal/content/104790). International travel is authorized for this contract. Travel is not to exceed $15,000.

Restricted or Limited Use of Data
The contractors may be required to sign and adhere to a non-disclosure agreement.

NMFS Project Contact
Michele Traver, NEFSC Acting Assessment Process Lead
Northeast Fisheries Science Center
166 Water Street, Woods Hole, MA 02543
Michele.Traver@noaa.gov Phone: 508-495-2195

Appendix 1. Red Hake Stock Structure Research Track Terms of Reference

1. Review and summarize all relevant literature on the existing stock structure of red hake in the northwest Atlantic.

2. Identify and evaluate any new and/or existing data relevant to the stock structure of red hake including but not limited to the species' life history (i.e. spawning, distribution, abundance, growth, maturity and natural mortality), morphometrics, and genetics.

3. Recommend the most likely biological stock structure among a set of alternatives from TOR2. Consider the current management unit as null hypothesis.

4. Evaluate existing experimental data on survey catchability of red hake. Examine the sufficiency of catchability data and, if appropriate, incorporate the catchability estimates into the assessment.

5. Apply the existing assessment model framework to the stock structure based on TOR 3 and 4 to ensure its utility in subsequent management track assessments. Evaluate existing reference points. Consider alternate assessment approaches if existing model framework does not perform well, and consider alternate reference points as needed.

6. Identify gaps in the existing research with respect to red hake stock structure. Develop a prioritized list of research recommendations to address these gaps. Comment on the feasibility and time horizon of the proposed research recommendations.

SAW Research Track TORs:

General Clarification of Terms that may be used in the Research Track Terms of Reference

Guidance to SAW Research Track Working Group about "Number of Models to include in the Peer Reviewer Report":

In general, for any TOR in which one or more models are explored by the Working Group, give a detailed presentation of the "best" model, including inputs, outputs, diagnostics of model adequacy, and sensitivity analyses that evaluate robustness of model results to the assumptions. In less detail, describe other models that were evaluated by the Working Group and explain their strengths, weaknesses and results in relation to the "best" model. If selection of a "best" model is not possible, present alternative models in detail, and summarize the relative utility each model, including a comparison of results. It should be highlighted whether any models represent a minority opinion.

On "Acceptable Biological Catch" (DOC Nat. Stand. Guidelines. Fed. Reg., v. 74, no. 11, 1-16-2009):

Acceptable biological catch (ABC) is a level of a stock or stock complex's annual catch that accounts for the scientific uncertainty in the estimate of Overfishing Limit (OFL) and any other scientific uncertainty…" *(p. 3208) [In other words, OFL ≥ ABC.]*

ABC for overfished stocks. For overfished stocks and stock complexes, a rebuilding ABC must be set to reflect the annual catch that is consistent with the schedule of fishing mortality rates in the rebuilding plan. *(p. 3209)*

NMFS expects that in most cases ABC will be reduced from OFL to reduce the probability that overfishing might occur in a year. (p. 3180)

ABC refers to a level of ''catch'' that is ''acceptable'' given the ''biological'' characteristics of the stock or stock complex. As such, Optimal Yield (OY) does not equate with ABC. The specification of OY is required to consider a variety of factors, including social and economic factors, and the protection of marine ecosystems, which are not part of the ABC concept. (p. 3189)

On "Vulnerability" (DOC Natl. Stand. Guidelines. Fed. Reg., v. 74, no. 11, 1-16-2009):

"Vulnerability. A stock's vulnerability is a combination of its productivity, which depends upon its life history characteristics, and its susceptibility to the fishery. Productivity refers to the capacity of the stock to produce Maximum Sustainable Yield (MSY) and to recover if the population is depleted, and susceptibility is the potential for the stock to be impacted by the fishery, which includes direct

captures, as well as indirect impacts to the fishery (e.g., loss of habitat quality)." (p. 3205)

Participation among members of a Research Track Working Group:

Anyone participating in SAW meetings that will be running or presenting results from an assessment model is expected to supply the source code, a compiled executable, an input file with the proposed configuration, and a detailed model description in advance of the model meeting. Source code for NOAA Toolbox programs is available on request. These measures allow transparency and a fair evaluation of differences that emerge between models.

Appendix 2. Draft Review Meeting Agenda

{Final Meeting agenda to be provided at time of award}

Red Hake Stock Structure Research Track Assessment

March 9-12, 2020

Stephen H. Clark Conference Room – Northeast Fisheries Science Center
Woods Hole, Massachusetts

DRAFT AGENDA* (version: December 3, 2019)

**All times are approximate, and may be changed at the discretion of the SARC chair. The meeting is open to the public; however, during the Report Writing sessions we ask that the public refrain from engaging in discussion with the SARC.*

Monday, March 9th, 2020

Time	Topic	Presenter(s)	Rapporteur
1:00 – 1:30pm	Welcome/Description of Review Process Introductions/Agenda/Conduct of Meeting	Michele Traver, Acting Assessment Lead TBD, Chair	
1:30 – 2:30pm	Review of Current Assessment and Historical Designations (TOR #1)	Toni Chute Dave Richardson, WG Chair	TBD
2:30 – 3:30pm	New Data and Analyses (TOR #2)	Dave Richardson, WG Chair	TBD
3:30 – 3:45pm	Break		
3:45 – 5:00pm	New Data and Analyses (TOR #2) cont.	Dave Richardson, WG Chair	TBD
5:00 – 5:30pm	Discussion/Review/Summary	Panel	TBD
5:30 – 5:45pm	Public Comment	Public	TBD
5:45pm	Adjourn		

Tuesday, March 10th, 2020

Time	Topic	Presenter(s)	Rapporteur
8:30 – 8:45am	Welcome/Logistics	Michele Traver,	

		Acting Assessment Lead TBD, Chair	
8:45 – 10:45am	New Data and Analyses (TOR #2) cont.	Dave Richardson, WG Chair	TBD
10:45 – 11:00am	Break		
11:00 – 12:30pm	Catchability (TOR #4)	Dave Richardson, WG Chair	TBD
12:30 – 1:30pm	Lunch		
1:30 – 3:30pm	Stock Structure Proposals (TOR #3)	Dave Richardson, WG Chair	TBD
3:30 – 3:45pm	Break		
3:45 - 5:00pm	Stock Structure Proposals (TOR #3) cont.	Dave Richardson, WG Chair	TBD
5:00 – 5:30pm	Discussion/Review/Summary	Panel	TBD
5:30 – 5:45pm	Public Comment	Public	TBD
5:45pm	Adjourn		
7:00pm	Dinner Social		

Wednesday, March 11th, 2020

Time	Topic	Presenter(s)	Rapporteur
8:30 – 8:45am	Welcome/Logistics	Michele Traver, Acting Assessment Lead TBD, Chair	
8:45 – 10:45am	Model Proposals (TOR #5)	Dave Richardson, WG Chair	TBD
10:45 – 11:00am	Break		
11:00 – 12:00pm	Research Recommendations (TOR #6)	Dave Richardson, WG Chair	TBD
12:00 – 12:30pm	Discussion/Review/Summary	Panel	TBD
12:30 – 12:45pm	Public Comment	Public	TBD
12:45 – 1:45pm	Lunch		

1:45 - 5:00pm	Peer Reviewer Summary Report Writing	Panel	
5:00pm	Adjourn		

Thursday, March 12th, 2020

Time	Topic	Presenter(s)	Rapporteur
9:00 – 5:00pm	Report Writing	Panel	

Appendix 3. Individual Independent Peer Reviewer Report Requirements

1. The independent Peer Reviewer report shall be prefaced with an Executive Summary providing a concise summary of whether they accept or reject the work that they reviewed, with an explanation of their decision (strengths, weaknesses of the analyses, etc.).

2. The report must contain a background section, description of the individual reviewers' roles 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 independent report shall be an independent peer review, and shall not simply repeat the contents of the Peer Reviewer Summary Report.

   a. Reviewers should describe in their own words the review activities completed during the panel review meeting, including a concise summary of whether they accept or reject the work that they reviewed, and explain their decisions (strengths, weaknesses of the analyses, etc.), conclusions, and recommendations.

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

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

   d. The report may include recommendations on how to improve future assessments.

3. The report shall include the following appendices:

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

Appendix 4. Peer Reviewer Summary Report Requirements

1. The main body of the report shall consist of an introduction prepared by the SARC chair that will include the background and a review of activities and comments on the appropriateness of the process in reaching the goals of the SARC. Following the introduction, for each assessment /research topic reviewed, the report should address whether or not each Term of Reference of the Research Track Working Group was completed successfully. For each Term of Reference, the Peer Reviewer Summary Report should state why that Term of Reference was or was not completed successfully.

   To make this determination, the SARC chair and reviewers should consider whether or not the work provides a scientifically credible basis for developing fishery management advice. If the reviewers and SARC chair do not reach an agreement on a Term of Reference, the report should explain why. It is permissible to express majority as well as minority opinions.

   The report may include recommendations on how to improve future assessments.

2. If any existing Biological Reference Points (BRPs) or BRP proxies are considered inappropriate, include recommendations and justification for alternatives. If such alternatives cannot be identified, then indicate that the existing BRPs or BRP proxies are the best available at this time.

3. The report shall also include the bibliography of all materials provided during the SAW, and relevant papers cited in the Peer Reviewer Summary Report, along with a copy of the CIE Performance Work Statement.

   The report shall also include as a separate appendix the assessment Terms of Reference used for the SAW, including any changes to the Terms of Reference or specific topics/issues directly related to the assessments and requiring Panel advice.