

**Center for Independent Experts (CIE) Peer Review Reports
on
Spiny Dogfish Research Track Assessment
and
Bluefish Research Track Assessment**

*by
Joseph Powers*

Executive Summary

Spiny Dogfish

The spiny dogfish assessment moved from a “stochastic estimator” model to an SS3 implementation of sex-specific length-based dynamics. A thorough evaluation of catch, indices, life history growth, and aging was conducted.

General lack of understanding of aging remains a weakness in the assessment. The SS3 configuration attempts to ameliorate the impacts somewhat, but it remains a problem. The model results show there are still inconsistencies. Nevertheless, it is recommended that the SS3 model be carried forward for management track evaluations. But it is recognized that some inconsistencies in SR model parameters, data weighting and uncertainties in catch will have to be reevaluated with the most recent data in the management track.

Bluefish

The bluefish research track assessment Working Group implemented a thorough examination of historical catches, indices, age-length keys, life history parameters. In so doing, it moved from the ASAP model to WHAM in a step-wise mode that was easy to follow and to justify.

There were a number of changes in data (most notably MRIPS recalibration), new and reevaluation of indices and updated natural mortality estimates, and the addition of selectivity blocks. The resulting WHAM model demonstrated minimal retrospective patterns in either F or SSB. Results of the base model are similar to the previous assessment but with some shifts in scale.

The WHAM model as configured should provide a basis for management track evaluations going forward and is, thus, recommended.

Background

Atlantic Spiny Dogfish (*Squalus acanthias*) and Atlantic bluefish (*Pomatomus saltatrix*) were selected for research track peer review in 2021. Two separate and unrelated Working Groups (WGs) were created to develop each research track.

The most recent full update stock assessment of the spiny dogfish (*Squalus acanthias*) was in 2018 and the method used was a Stochastic Estimator approach. Based on that assessment, spiny dogfish was not overfished and overfishing was not occurring, although there were observed survey SSB decreases in recent years, especially in 2021. The 2022 research track spiny dogfish assessment WG moved the base model to Stock Synthesis (SS3).

The most recent full update stock assessment for bluefish (*Pomatomus saltatrix*) was in 2021, and it was based on an ASAP (age-structured assessment program ASAP) model peer-reviewed in the 2015 benchmark assessment. The MRIP (Marine Recreational Information Program) calibration resulted in an increase in the estimated recreational catch and caused scale changes in both biomass and reference points. The bluefish population was overfished, and overfishing was not occurring according to the 2021 assessment. The 2022 research track bluefish assessment moved from ASAP to a Woods Hole Assessment Model (WHAM).

These two Research Track assessments allow for evaluating and using new datasets, models, or stock structures. The stock assessments are expected to provide the basis for future management track assessments. To that end, the separate WG efforts were evaluated through a single Spiny Dogfish and Bluefish Research Track Stock Assessment Peer Review Panel that met via WebEx from December 5 through December 9, 2022. The Panel was composed of three scientists selected by the Center for Independent Experts (CIE): Paul Medley, Robin Cook, and Joseph Powers (me). Additionally, the Panel consisted of the Chair: Yan Jiao (member of the New England Fisheries Management Council Scientific and Statistical Committee). The review focused on 18 Terms of Reference (TORs, 9 for each species), listed below). The Panel addressed each of those TORs in turn, based on the WG's documentation and the feedback and opinions of the contributing scientists.

Description of the Individual Reviewers' Roles in the Review Activities

The role of each of the Center of Independent Experts (CIE) in this review was to attend and participate in the panel review meeting, conduct independent peer reviews in accordance with the requirements specified in this Performance Work Statement (Appendix 2) and TORs (Appendix 2, Annex 1), to assist the Peer Review Panel Chair with contributions to the Peer Reviewer Summary Report and to deliver individual Independent Reviewer Reports to the CIE accordingly explaining whether each research track Term of Reference was or was not completed successfully.

In particular my responsibility as a CIE reviewer is to deliver an independent report addressing the TORs in the work statement. This document represents my report. My specific independent responses to each of the 18 TORs and my overall conclusions for the two reviews follow.

Summary of Findings for Each TOR

Spiny Dogfish

Generic Research Track TOR 1. *Identify relevant ecosystem and climate influences on the stock. Characterize the uncertainty in the relevant sources of data and their link to stock dynamics. Consider findings, as appropriate, in addressing other TORs. Report how the findings were considered under the TORs.*

This TOR has been met.

Several hypotheses of ecosystem/climate influences on spiny dogfish were explored by the WG including possible shifts in spatial distribution (VAST models, vessel trip reports), but nothing substantial was shown.

Life history information was reevaluated (growth, maturity) and it was shown that recent length at 50% maturity was significantly less than observed in studies prior to 2010. The new ageing data were rejected by the WG because of the high measurement error. Nevertheless, the spine ageing analysis and the tag-recapture analysis suggested a decreased growth curve after mid-2005 compared with the spiny dogfish sampled in the 1980s. This led to alternative configurations of the biological models in SS3 which I support.

SR with environmental factors did not improve model fitting however these were based on the swept area version of the SR data and not estimated in SS3. Additionally, as noted in my response to Spiny dogfish TOR4, the application of the SR model in SS3 needs to be revisited.

Generic Research Track TOR 2. *Estimate catch from all sources including landings and discards. Describe the spatial and temporal distribution of landings, discards, and fishing effort. Characterize the uncertainty in these sources of data.*

This TOR has been met.

I have limited comments on the landings and discard data from the fishery sectors, as the data were collected with standard methodologies. In converting to length and sex compositions, there were gaps that were filled by “pooling strategies”. These are never entirely satisfactory, but are the best way forward at this time. The landings and discards uncertainty were quantified and reported in the assessment report. However, these uncertainties were not accounted for in the SS3 model runs because of convergence issues. That uncertainty should be considered in future SS3 model configurations under TOR4.

Generic Research Track TOR 3. *Present the survey data used in the assessment (e.g., indices of relative or absolute abundance, recruitment, state surveys, age-length data, application of catchability and calibration studies, etc.) and provide a rationale for which data are used. Describe the spatial and temporal distribution of the data. Characterize the uncertainty in these sources of data.*

This TOR has been met.

The WG examined many options for fishery independent survey data (9 surveys were considered). Additionally, VAST methods were used to explore the combination of indices. In the final analysis only the NEFSC bottom trawl survey index was recommended in the SS3 base model run while the others were considered in sensitivity runs. The NEFSC bottom trawl survey index was the most spatially broad. However, even this index was not a strong indicator of trend. In general, the potential influence of indices was hard to assess because of no data weighting and poor fits to the indices. Because of the importance of estimating cohort signals and given the pup/recruitment data seen in the NEFSC spring bottom trawl survey length frequency, which was not well explained in the SS3 model fit, I recommended that a pup index may be considered in the future as a recruitment index, but with special attention to the timing of the index within the year and how that relates to the Stock Recruitment relationship.

Generic Research Track TOR 4. *Use appropriate assessment approach to estimate annual fishing mortality, recruitment and stock biomass (both total and spawning stock) for the time series, and estimate their uncertainty. Compare the time series of these estimates with those from the previously accepted assessment(s). Evaluate a suite of model fit diagnostics (e.g., residual patterns, sensitivity analyses, retrospective patterns), and (a) comment on likely causes of problematic issues, and (b), if possible and appropriate, account for those issues when providing scientific advice and evaluate the consequences of any correction(s) applied.*

This TOR has been met, but there are some reservations.

The SDWG chose to move to the SS3 framework for this research track assessment primarily because it allowed for sex-specific analyses. The underlying population model was fitted to length data. Attempts were made to utilize age-reading data from the recent period. However, problems in the age-reading, especially for the younger fish led to an overall rejection of the data. But time blocks were constructed for biological parameters in the SS3 configuration. These blocks, along with the catch data, indices/fisheries and selectivities, and a stock-recruitment model were integrated into the SS3 framework.

There were some concerns about the SDWG recommended SS3 base model run because of the lack of meaningful model comparisons among sensitivity runs and the model run itself because of data weighting and fixed parameters.

The SS3 configuration was based on a suite of “fleets” comprised of two landings fleets, three discarding fleets and one survey fleet: the NEFSC spring bottom trawl survey. The landings and discards from the same fishery were separated into two fleets. The base run only included the NEFSC spring bottom trawl index and the model did not fit the survey index well (shown as a flat line that did not capture the historical decrease well). The base model only down-weighted the length frequency of the NEFSC survey fleet but not the landings and discards fleets, which was probably the reason for the poor fit to the survey index. The model did not capture the recruitment signal well, which was mainly from the NEFSC survey, and the fit to the small-sized group length compositions was consistently poor. Seventeen sensitivity runs tested the influence of the growth model setup, the mortality assumption, the SR relationships, the biology time blocks, the NEFSC bottom trawl selectivity time blocks, the model starting year, and the use of the survey indices. The sensitivity runs did not re-weight the data, so none fitted the survey indices well, and it was hard to diagnose the influence of each model change. According to the SDWG, many model runs did not converge if the data weighting was turned on or manually fixed, which made diagnostics on alternative model runs difficult to interpret during the research track review.

Spiny Dogfish Stock-Recruitment

Additionally, the base model generated stock-recruitment data in the early years that appeared anomalous. For those reasons I have gone into some depth about the stock-recruitment relationship, its theoretical underpinnings and how it is employed in this instance. The stock-recruitment function used was the Taylor et al. function: *Taylor, Gertseva, Methot, Maunder 2013. A stock-recruitment relationship based on pre-recruit survival, illustrated with application to spiny dogfish shark. Fisheries Research 142: 15-21.* Note there is a typo in the function equation in the assessment report and the assessment presentation, but that typo was not carried into their analysis (i.e., a typo, not an analysis mistake!). In what follows, I develop the background of this function and then show how that relates to the spiny dogfish application and fitting. As will be seen, the selection of dogfish recruitment parameters is highly related to important characteristics of spiny dogfish life history. For typical groundfish, these parameters are much more flexible, and we typically don’t worry too much about their scale, but with dogfish, we do.

Going to the basics of Taylor et al.:

$S_0 = R_0/B_0$ = the equilibrium survival rate with no fishing at the beginning of the year
 R_{max}/B_{max} = the equilibrium maximum survival rate at the beginning of the year (maximum slope at the origin)
 R_d/B_d = the equilibrium maximum survival rate at some time dt within ages 0 to 1
 $= (R_{max}/B_{max}) \exp(-m)$ where density dependence starts

$$z_{frac} = - \frac{\ln(R_d/B_d) - \ln(R_0/B_0)}{\ln(R_0/B_0)} = 1 - \frac{\ln(R_d/B_d)}{\ln(R_0/B_0)}$$

$$(1 - z_{frac}) \ln(R_0/B_0) = \ln(R_d/B_d)$$

$$\frac{R_d}{B_d} = \left(\frac{R_0}{B_0}\right)^{1-z_{frac}}$$

$$R_y = B_y \exp \left(\ln \left(\frac{R_0}{B_0} \right) \left\{ 1 - z_{frac} \left[1 - \left(\frac{B_y}{B_0} \right)^\beta \right] \right\} \right)$$

$$\text{when } B_y = B_0 \text{ then } \frac{R_y}{B_y} = \frac{R_0}{B_0}$$

when $B_y \rightarrow 0$ then $\frac{R_y}{B_y} \rightarrow \left(\frac{R_0}{B_0}\right)^{1-z_{frac}}$ the slope at the origin at beginning of the year

Thus,

$$R_y = \frac{R_d}{B_d} B_y \exp \left(\ln \left(\frac{R_0}{B_0} \right)^{z_{frac}} \left(\frac{B_y}{B_0} \right)^\beta \right)$$

$$R_y = a B_y \exp \left(-b B_y^\beta \right)$$

$$b = \ln \left(\frac{R_0}{B_0} \right)^{\frac{z_{frac}}{B_0^\beta}}$$

$$a = \frac{R_d}{B_d} = \left(\frac{R_0}{B_0}\right)^{1-z_{frac}}$$

So basically the Taylor et al function is a Ricker function with the added parameter beta.

The Taylor et al model can also be expressed as a two-step recruitment process:

$$\frac{dR_t}{dt} = -MR_t \quad \text{for } 0 < t < \Delta t \quad \text{and}$$

$$\frac{dR_t}{dt} = -GR_t - D R_{t+\Delta t}^\beta \quad \text{for } \Delta t < t < 1$$

The solution (Brooks and Powers 2007) (letting $R_{t=0}$ =age 0 pups at beginning of the year) is

$$R_{t+\Delta t} = R_{t=0} \exp(-M\Delta t)$$

$$R_{t+1} = R_{t=0} \exp(-M\Delta t) \exp[-G(1-\Delta t)] \exp \left\{ -D \left(R_{t=0} \exp(-M\Delta t) \right)^\beta (1-\Delta t) \right\}$$

$$R_{t+1} = a B_t \exp \left(-b B_t^\beta \right)$$

$$a = \exp(-G(1-\Delta t) - M\Delta t)$$

$$b = D \left(\exp(-M\Delta t) \right)^\beta (1-\Delta t)$$

which is the same as above and the same as Taylor et al.

So we have modified-Ricker, Ricker, and Beverton-Holt:

$$R_{t+1} = aB_t \exp(-bB_t^\beta)$$

$$R_{t+1} = aB_t \exp(-bB_t)$$

$$R_{t+1} = \frac{aB_t}{1 + bB_t}$$

Note, in all three cases the “a” value can be parameterized as $a = \left(\frac{R_0}{B_0}\right)^{1-z_{frac}}$

So the underlying assumption of a Ricker-like construct (which Taylor et al. is) is that all the density-dependence occurs instantaneously at the beginning of the Ricker-like process which in this case is at Δt . There is no continuous density-dependence acting throughout the interval. The density-dependent action in a Ricker-like construct would correspond to a major short-term transition at which density mortality acts over a short period of time, such as birth of pups, transition to a new habitat, forming into schools and matching eggs with sperm to name a few. In this case it would seem natural to model Δt as being the time within the year when pups are born. But in the SS3 construct apparently this can't be done. So, in order to normalize the pups born at $t = \Delta t$ back to the pups in the womb at $t=0$, the method inflates the number of pups by a mortality rate. This mortality rate relates to both the mortality rate of pups in the womb and the natural mortality rate of pregnant females in the second year of pregnancy during the first part of the year Δt ; **AND the fishing mortality rate during the first part of the year on that subset of adult females**. However, that embryo/pregnant female fishing rate has been ignored in this spiny dogfish application, or more specifically it is assumed to be either small and/or constant from year to year. The parameter z_{frac} encompasses the “in-the-womb” mortality rate, so the choice of z_{frac} is implicitly making the F of embryos constant assumption.

Before going further, let me make a few comments on the Taylor et al. (2013) paper from whence this stock recruitment (SR) relationship came, the integration into SS3 and the use of “steepness”. First, all SRs are survival functions $R/B=a \times c/f(B)$ where c is the number of eggs/weight of females. Where fecundity data were available, these have been used in many stock assessments many, many years prior to 2013. Thus, there was nothing unique about using a “survival function” implied by Taylor et al. Secondly, as shown above, Taylor et al. is a modified Ricker function which has been hypothesized for years (I believe even by Ricker himself). What is unique in Taylor et al. is the reparameterization into a form that assists selecting a prior on the slope at the origin (although I personally find the z terminology much less intuitive than the (R/B) terminology given above). Further, the Ricker and BH models can also be parameterized using either the z terminology or the (R/B) terminology. The Taylor et al. paper doesn't provide much discussion of why an additional parameter Beta is justified. The BH and Ricker require R_0 to scale the process, (B_0/R_0) from life history data and steepness, the modified Ricker uses R_0 ,

(B_0/R_0) and z_{frac} , but then adds a beta. I suppose since z_{frac} is fixed in most applications, beta is the only parameter that gives the model some flexibility.

The key parameter in all stock recruitment models is the slope at the origin which defines the theoretical maximum productivity. The use of the steepness parameterization in the BH model has been useful because it provides some intuitive visualization of what the slope at the origin is. While Ricker and modified Ricker can be parameterized in terms of steepness, steepness in those models doesn't convey much information about the slope at the origin. And comparisons of steepness between Ricker, modified-Ricker and BH doesn't tell me much until I transform back into the model parameters. But I gather that SS3 is structured around using steepness, so there we are.

Both Taylor et al. and this assessment report notes that SS3 is structured such that B is measured at the beginning of a year interval (January 1 in this case) and recruits are measured at the end of the elapsed year (Dec 31) and that there is no fishing mortality during that year. The Taylor et al. paper goes through an elaborate parameterization to adjust to the rigidity of SS3. I guess they felt that this was better than doing the work of coding in SS3 to make the timing of the stock recruitment process during year 0 more flexible, even though one of the authors was the main developer of SS3. But it is now a decade later and as I noted, the model above ignores the dynamics of fishing on late-pregnancy females because of the contrived process to adjust to SS3's rigidity.

But now let's return to this spiny dogfish application. The base model used was $\beta=1.5$; $z_{frac}=0.9$; $B_0=7.15$, $R_0=1.6$ (the latter two by eyeballing from graph) giving $R_0/B_0=0.224$ which implies $B_0/R_0=4.47$ pups per recruit. The slope at the origin $= (R_0/B_0)^{(1-z_{frac})}=0.86$ which implies an "effective mortality rate" of $-\ln(0.86)=0.15$. Also, σ_R was set at 0.3 without much justification that I could see.

I plotted the %SPR diagonals in the following and overlaid them on the model-generated stock-recruitment data (Figure 1).

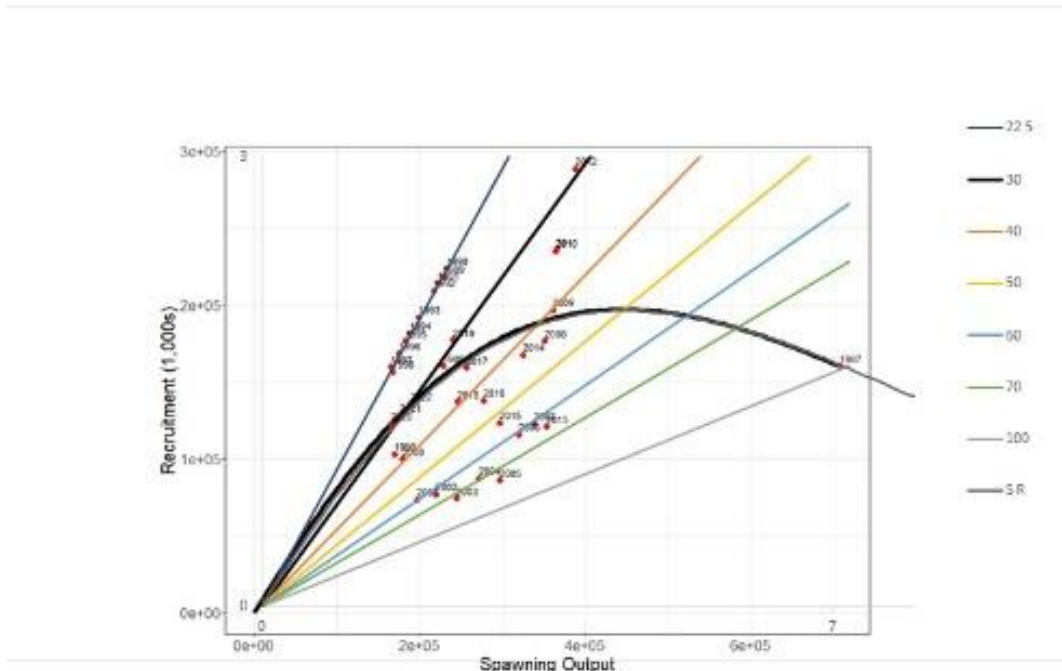


Figure 1: Spawning output based on recruitment

A couple of key things to note: 1) the slope at the origin 0.86 implies an SPR at the origin of about 22.5%; 2) the early years data lines up exactly with the slope at the origin. I suspect that SS3 has imposed a constraint/penalty that the maximum R/B cannot exceed the slope at the origin. The model wants to generate higher recruits to accommodate the shift in Linf. But these have been artificially constrained in the base runs. This needs to be rectified when using the model in actual application in the future. However, a mitigation is the result in the assessment report Fig 4.4.3 where the early data were discarded (and thus the constraint did not need to be imposed). The late year trends were very similar with a small change of scale in that case. I suspect this problem was caused by a σ_R of 0.3 which is much larger than the effective mortality rate at the origin of 0.15. The way to explain this is that σ_R is slapping the $\exp(\sigma)$ onto the numerator of the stock-recruitment function which effectively is assuming that this affects the density independent (di) parameter but not the density dependent (dd) one. For most stocks of fish this is not a big deal because the slope at the origin is very low. But for sharks it is a big deal. The density-independent rate is pretty constrained so σ_R on that rate should be lower. Ideally, one could have different sigmas on the dd and di rates. I believe WHAM can do this. But I doubt if SS3 can. Another alternative is to not assume log-normal σ_R on Recruits, but rather to model survival $[(R_t/B_t) \text{ given } B_t]$ as, perhaps, a Beta pdf. However, the latter is a longer term activity addressed under TOR 9.

By developing the Taylor et al. model from first principles, I have shown that it can be couched as two sequential processes and that the timing of each changes the effect of fishing on the contribution by late pregnant females to recruitment. Since in the final model the timing issues are not specified directly in the collapsed parameters, there are different ways to interpret the underlying process that has been used. In particular, we can avoid the effect of late pregnancy

fishing by assuming that the second stage starts very soon after January 1 (therefore no fishing mortality) and that the density-dependent effect occurs instantaneously after that. This implies that a density dependent event occurs in the womb on January 2. Not much justification there, but that is the interpretation with no fishing. Alternatively, one could assume that there is a constant fishing rate in some portion of the recruitment age over all years at all stock sizes that is incorporated into zfrac. But then that rate would still be in the zfrac when there is no fishing, implying that B0 is actually larger than what is being estimated. Not much justification there either. And of course there is a continuum between these two interpretations.

Also, note that in this application the choice of zfrac fixes the slope at the origin:

$a = \left(\frac{R_0}{B_0}\right)^{1-zfrac}$. A fairer comparison of sensitivities with Ricker and BH would be to fix their slopes at the origin to the same value. The report says that their steepness values were fixed but not that the fixed values gave the same slope at the origin.

Finally, the choice of zfrac, itself. This particular choice of 0.9 implies that when the stock size is near zero, the effective mortality (EM) rate in the first year is 0.15. Then when the stock is near B0 the effective rate = $-\ln(R_0/B_0)=1.5$. This is compared to the M value at age 1 of around 0.3 (as memory serves). So recruits from a severely depleted stock live the good life of an EM of 0.15 and then are faced with the holy-hell of age 1 with M=0.3. To some extent this is convoluted with the above interpretation of timing of density dependence and the protection provided by pregnancy. But it is food for thought. One might suggest the EM at the origin can't be lower than age 1 M. This has a certain logic, but I can't vouch for how it would interact with actual data. I recall that Pacific coast thresher shark used a zfrac value of 0.6, but I do not remember the life history that led to this.

I am not convinced that a BH or Ricker with the same slope at the origin wouldn't perform almost as well (using one fewer parameter) than the modified Ricker. One could try a null model (BH with steepness of one), but there would have to be a penalty that R/B could never be greater than 1. But in effect that would not be all that much different than a BH with a fixed slope at the origin. Also, use of Ricker or BH still doesn't get around the problem of assuming spawning output occurs on January 1 and not sometime mid-year (the SS3 rigidity effect).

So even after this long discourse on stock-recruitment relationships and the difficulties of the present application, I can still support the SS3 SR implementation going forward towards a management track assessment. My reasons include the results in Fig. 4.4.3; the acceptance of SPR 60% in the Reference Point (TOR 5) and projections (TOR 6) and the comparison with the alternatives (TOR 8). But as with any application, the fixed parameter values will have to be re-examined based on the incremented data set and interpreted along the lines of that given above to assure some consistency. Also, the SS3 constraint/penalty and sigmaR conflict that causes the early R/B data to equal the slope at the origin must be rectified.

Generic Research Track TOR 5. Update or redefine status determination criteria (SDC; point estimates or proxies for BMSY, BTHRESHOLD, FMSY and MSY reference points) and provide estimates of those criteria and their uncertainty, along with a description of the sources of uncertainty. If analytic model-based estimates are unavailable, consider recommending alternative measurable proxies for reference points. Compare estimates of current stock size and fishing mortality to existing, and any redefined, SDCs.

I compared the equilibrium results of the reproductive output** per recruit analyses from the assessment report to the equilibrium results from the stock recruitment relationship (eyeballing from my Figure 1 above):

	SPR50%	SPR60%	SPR70%
Eq. Spawning Output	268,707	370,799	457,116
F	0.037	0.025	0.017
Eq. Catch (MSY proxy)	18,876	16,792	12,657
Eq. Spawning Output	450,000	530,000	580,000
From SR model			

**To us fishery assessment types, use of terminology like spawning output is fine, as we all know what it really means. But to a shark biologist, they will probably have conceptions about our lack of knowledge of shark life history, thinking that sharks actually spawn. In a previous life someone once called me to task on this and I never forgot.

Also, by eyeballing from the graph, it is implied that the average recruitment used to scale the per recruit results was 20-30% lower than equilibrium R0 from the SR relationship. Also, note that from the per recruit analyses and the graph, $\frac{1}{2}BSPR50\% \sim 135,000 \sim BSPR30\%$, $\frac{1}{2}BSPR60\% \sim 185,000 \sim BSPR35\%$ and $\frac{1}{2}BSPR70\% \sim 230,000 \sim BSPR45\%$. If the definition of an overfished stock is $\frac{1}{2}$ of the BMSY surrogate, then the $\frac{1}{2}BSPR60\%$ result is relatively close to the selected slope at the origin. Thus, the overfished criterion as being $\frac{1}{2}$ of the BMSY surrogate may have to be revisited with spiny dogfish. However, this relates to my discussion about Zfrac, the slope at the origin, the selected SR relationship and sigmaR under TOR 4. All of these are related.

There are long term implications of using the SR results vs the average R approach to the per recruit analyses as exemplified by the above table. But all in all, **I agree that this TOR has been met and I agree with using SPR60% as a proxy for BMSY going into the next management track assessment.** But my reasons differ somewhat from the justification of the WG. Despite the mismatch of the equilibrium reference points implied by the SR relationship with those of the per recruit analyses, I support the per recruit approach. I tend to discount some of the recruitment results from the early years due to SR problems (TOR 4), and thus, my SPR60% choice is based on life history, current recruitment values and the recognition that neither choice provides much insight into long term productivity.

But care needs to be taken in determining consistency of parameters/models in the final management track results. That consistency is important given the low productivity of spiny dogfish.

Generic Research Track TOR 6. *Define appropriate methods for producing projections; provide justification for assumptions of fishery selectivity, weights at age, maturity, and recruitment; and comment on the reliability of resulting projections considering the effects of uncertainty and sensitivity to projection assumptions.*

This TOR has been met with some concerns.

The WG conducted 3-year (2022-2024) short-term projections under $F=0$, $FSPR_{70\%}$, $FSPR_{60\%}$, and $FSPR_{50\%}$ using the SS3 internal projection tool processes and uncertainty in recruitment and numbers-at-age. Fleet selectivity, maturity, natural mortality, the SR relationship, and growth are the same as estimated from the 2012-2019 period from the SS3 model run. The 3-year projection showed a sharp decrease in 2020 but increased after that, likely due to the maturation of many females in the large 2009-2012 year classes. There are concerns related to the definitions of fleets in SS3. These are in effect “pseudo fleets” that separate catch components into landings and discard “fleets” while combining gears in different groups. As a result, forward projections with different F multipliers assume particular fleet selectivity and discard selection coefficients that may be unrealistic. Furthermore, given the artificial nature of the model fleets, it is unclear how these relate to management, making the interpretation of potential interventions problematic. This is a standard problem in projections. %SPR always is affected by the relative selectivities between fleets. Thus, alternative regulatory schemes will alter selectivities and shift the F-multipliers. At the early stages of doing projections, these current methods are adequate. But as the regulatory scheme evolves, the projections may have to be revisited. The WG should be aware of this when communicating results.

The NEFSC bottom trawl survey swept-area estimated SSB_{2021} indicated a large decline; the projection did not capture this decline. Noting the concerns from the SS3 model runs not fitting the survey abundance index (indices), future diagnostics on both the assessment model and projection should be evaluated between the research track review and management track review. Consistency is important with the low productivity spiny dogfish in the SR parameters and likely short-term projections.

Generic Research Track TOR 7. *Review, evaluate, and report on the status of research recommendations from the last assessment peer review, including recommendations provided by the prior assessment working group, peer review panel, and SSC. Identify new recommendations for future research, data collection, and assessment methodology. If any ecosystem influences from TOR 2 could not be considered quantitatively under that or other TORs, describe next steps for development, testing, and review of quantitative relationships and how they could best inform assessments. Prioritize research recommendations.*

This TOR has been met.

SDWG has made substantial progress on several of the research recommendations stemming from the 43rd SAW Stock Assessment Report (NEFSC 2006), MAFMC 2020-2024 Research

Priorities (2019), and MAFMC SSC Research Recommendations in 2019. A great deal of progress was made, the results of which were incorporated into the research track assessment model. The SDWG also developed new research recommendations, but thought the first one below is the most important:

- Consistently collect, process, and age spines of spiny dogfish to understand growth and support future age-based assessments.
- Continue exploration into the spatial distribution of spiny dogfish (e.g., off-shelf abundance).
- Further explore the sensitivity of the SS3 model parameterization and configuration.
- Conduct directed studies that estimate discard mortality rates for spiny dogfish by commercial and recreational harvesting gear type.
- Develop state-space models that can fit to length data.
- Investigate drivers in the decline in maturity over time.
- Continue developing the VAST models presented.
- Investigate datasets enumerating the abundance or diet of known spiny dogfish predators for insight into natural mortality rates.

I agree with the general tone of the new research recommendations, particularly the collection and analysis of aging data. In addition, I recommend further exploration into the use of other survey abundance indices and fishery catch rate that may inform either YOY or large spiny dogfish and how these relative to the SR models being used.

Generic Research Track TOR 8. *Develop a backup assessment approach to providing scientific advice to managers if the proposed assessment approach does not pass peer review or the approved approach is rejected in a future management track assessment.*

This TOR has been met.

The spiny dogfish assessment method used prior to this research track has been the “stochastic estimator”. Thus, by default one would have reapplied this method if the research track assessment had not been done. This method would also be used if SS3 approaches failed, and an alternative was not shown to be better. To that end, the WG explored three other alternatives (Ismooth, Average Catch and Depletion-based Stock Reduction Analysis). The WG concluded that these analyses were unlikely to perform better than the stochastic estimator. I agree that under these circumstances the stochastic estimator was appropriate as a backup method. However, it is expected that the SS3 model framework will be the appropriate way forward for the management track assessment (subject to further consideration on data weighting, sensitivity analyses, and parameter selection mentioned in the other TORs).

Generic Research Track TOR 9. *Identify and consider any additional stock specific analyses or investigations that are critical for this assessment and warrant peer review, and develop additional TOR(s)* to address as needed.*

Note that this TOR was not within the WG's mandate and was not addressed explicitly by the WG or in the review meeting. However, it is still within the CIE's PWS for this review. Thus, I have used this TOR 9 as an opportunity to **identify and consider** possible avenues for more generic assessment research that would assist not only future dogfish assessments, but other stocks as well.

It seems that shark stock assessments are typically driven by low-productivity SR relationships. But those SR models are not well depicted in the standard assessment framework such as SS3 where the timing of when the recruitment process starts, when it ends and what constitutes a recruit are often fixed under a typical groundfish narrative. It would seem to me that a more flexible shark SR construct would numerically integrate over the relevant recruitment time span. The sacrifice of computation time to do numerical integration versus a standard one period BH or Ricker or modified Ricker is not likely to be excessive.

Additionally, I would recommend the exploration of using a beta model of survival rather than a lognormal sigmaR in generating deviates, per my discussion in TOR4.

Atlantic Bluefish

Generic Research Track TOR 1. Identify relevant ecosystem and climate influences on the stock. Characterize the uncertainty in the relevant sources of data and their link to stock dynamics. Consider findings, as appropriate, in addressing other TORs. Report how the findings were considered under the TORs.

This TOR has been met.

The WG reviewed the existing research and synthesized the existing data on the social-economic, ecosystem, and life history. The WG developed new analyses, including VAST species distribution changes, ecosystem indicators, and a VAST forage fish index, and applied their findings to the stock assessment model runs. The findings reconfirmed that Gulf of Mexico catch data should not be included in this assessment. This is consistent with what was done in the past.

Seasonal length frequencies and length-weight relationships at a minimum were used, and a seasonal-regional level of data was used where possible. The WG also developed age-specific natural mortality, which was used in the recommended WHAM model run for management purposes. The WG also addressed several previous research recommendations on life history, species distribution, and recruitment with environmental factors.

Although the forage fish index was not used in the recommended WHAM model run, its influence was tested in one model run as a covariate of MRIP catchability and was suggested for further research. The review panel also suggested that it may be considered in a catch rate

standardization step before being used in the WHAM model to better understand the catchability changes of the MRIP CPUE.

A great deal of work was done to address this TOR and the work was well done.

Generic Research Track TOR 2. Estimate catch from all sources including landings and discards. Describe the spatial and temporal distribution of landings, discards, and fishing effort. Characterize the uncertainty in these sources of data.

This TOR has been met.

Commercial and recreational landing and discard total number, total weight, length distribution, and release mortality were all updated and, in several cases, significant changes were made.

The commercial discards had not been estimated in the past but were included in this assessment. This was incorporated with a 32% release mortality based on literature review. The recreational release mortality was updated from 15% to 9.4% based on literature reviews, including the most recent research.

Of particular importance was the recalibration of the time series of recreational catch based on new methods for effort estimation (FES) and on intercepts of the survey (AP AIS). These new methods evolved to address the outmoded telephone survey and the allocation of intercept sampling. While the new and old methods were run jointly for several years, models to recalibrate the historical data accordingly were required. Thus, the scale of the recreational catches in general for all species was changed and bluefish was no exception.

Recreational length frequency was calculated by accounting for the differences among seasons and regions, which was further used in developing seasonal catch-at-age to account for fish size variation among seasons and regions.

In general, the methodologies were consistent with other stocks and the general availability of data. And, as noted, considerable effort was made on recreational MRIP catch data.

Discard mortality will always be a result of the general practices within a fishery. Thus, it should always be monitored to the extent feasible. In particular, the use of alternative hooks and size selectivity of gear might be addressed in the future.

Generic Research Track TOR 3. Present the survey data used in the assessment (e.g., indices of relative or absolute abundance, recruitment, state surveys, age-length data, application of catchability and calibration studies, etc.) and provide a rationale for which data are used. Describe the spatial and temporal distribution of the data. Characterize the uncertainty in these sources of data.

This TOR had been met.

Several new research items were included for developing relative abundance indices and age length keys (ALK). A Bayesian hierarchical model was used following Conn (2010) to develop a composite YOY index instead of using 6 separate YOY indices.

All the survey indices, except the SEAMAP indices of ages 0 and 1, were developed using both a design-based approach and a model-based approach (GLM framework used), and the WG decided to use the results from the design-based approach. The trends of the survey indices are consistent between the two methods and which method was used did not influence the output much according to the corresponding sensitivity runs. The WG felt that by using the design-based approach it would be easier to maintain consistency for future updates. I agree.

An important improvement in the MRIP CPUE was the use of a guild approach to select trips that were purportedly “targeting” bluefish. The “guild” was a suite of like species that fishers often encounter. By filtering the trips by this criterion, the resultant CPUE more faithfully depicts trends in abundance and reduces the uncertainty in catchability. Hence a better CPUE index. The definition of targeted trips has long been a weakness in recreational CPUE analyses, and this is a good improvement.

A multinomial model was used to estimate probabilities in the age-length key, which avoided having to use an ad hoc “borrowing” method for empty cells when the sample sizes were small. The method seems reasonable for bluefish in this case. Multinomial ALKs used to derive age composition data may have the effect of implying these data are more precise than is actually the case. A potential issue with pre-processing the data in this way may be to give more weight to the composition data relative to the abundance data. A sensitivity run using ALKs applying the older “borrowing” method resulted in poorer model convergence but lower F and higher SSB. So, this might be revisited in the future.

Generic Research Track TOR 4. Use appropriate assessment approach to estimate annual fishing mortality, recruitment and stock biomass (both total and spawning stock) for the time series, and estimate their uncertainty. Compare the time series of these estimates with those from the previously accepted assessment(s). Evaluate a suite of model fit diagnostics (e.g., residual patterns, sensitivity analyses, retrospective patterns), and (a) comment on likely causes of problematic issues, and (b), if possible and appropriate, account for those issues when providing scientific advice and evaluate the consequences of any correction(s) applied.

This TOR had been met.

The last benchmark stock assessment (2015) used an ASAP model and has been used for management purposes since then. It was last updated in 2021 in a management track assessment. The WG moved the assessment from ASAP to a WHAM framework in this research track assessment review. The stepwise migration from ASAP to WHAM, with the inclusion of

new data or parameter configuration, was clear, well thought out, and reasonable. The WG included a continuity ASAP run followed by a bridge model built with new data, smoothed age-length keys, age dependent M, new selectivity blocks, and other parameter configurations. This was then moved to the new WHAM model framework. This step-by-step approach, including results and diagnostics, helped the review process.

I am especially enamored with the integration of estimated process error in the WHAM framework. Essentially, in the WHAM framework the underlying biology/ecology is being modeled. The estimated variability determines the scope of response that a stock is likely to demonstrate. The inherent variability serves to define the likely cumulative effects of suites of climate and environmental shifts without focusing on individual indices whose signals may vary from year to year.

To that end, the model estimates include process error on the numbers at age in the model and the use of fixed natural mortality varying with weight (Lorenzen M). I am also, especially, supportive of the MRIP guild approach to define targeting in the recreational CPUE. This is a good step forward for the use of recreational data. The base WHAM model was appropriate for management purposes to provide scientific advice.

Previous assessments were heavily dependent on the MRIP CPUE. The new assessment is no longer as reliant on this index. The assessment is, however, now more sensitive to the PSIGNS index, the removal of which results in lower SSB and higher F. The PSIGNS index contained most of the information on the older fish abundance. However, given that this survey is limited in geographical coverage, some care is merited in interpreting the results.

The WG investigated the use of a forage index to account for changes in survey catchability. This was an innovative approach that merits further analysis.

The WHAM framework-based model may further consider processes such as natural mortality and fish spatial distribution changes based on what the WG found in TOR1.

Generic Research Track TOR 5. Update or redefine status determination criteria (SDC; point estimates or proxies for BMSY, BTHRESHOLD, FMSY and MSY reference points) and provide estimates of those criteria and their uncertainty, along with a description of the sources of uncertainty. If analytic model-based estimates are unavailable, consider recommending alternative measurable proxies for reference points. Compare estimates of current stock size and fishing mortality to existing, and any redefined, SDCs.

This TOR had been met.

The WG continued the use of SSB_{35%} as the SSB_{MSY} proxy and used the last five-year average Weight at Age (WAA) and selectivity for reference points estimation. The WG agreed that the literature (Rothschild et al. 2012; Thorson et al. 2012) supported the use of F_{35%} for bluefish and continued the use of F_{35%} as the F_{MSY} proxy. It was acknowledged that it was the generally

accepted approach in this region to use SPR analysis for the reference points estimation. Both $F_{35\%}$ and $SSB_{35\%}$ were calculated internally in WHAM using average recruitment over the time series (1985-2021), and 5-year averages for fishery selectivity and weights-at-age for SSB per recruit calculations $F_{35\%} = 0.248$. $SSB_{35\%}$ was calculated using SPR at 35% (0.718), and the mean of the full time series of recruitment (127,924 tonnes) $SSB_{35\%} = 91,897$ tonnes. Natural mortality and maturity were assumed constant over time in the model. Uncertainties of the BRP estimations were included in the assessment report, shown as Cis, and were calculated internally in WHAM and the marginals of the Kobe plots are quite useful.

The review panel discussed whether $SPR_{35\%}$ was the best proxy of MSY. Based on the plot of $YPR|F$ and $SPR|F$, the $SPR_{35\%}$ is less than F_{max} but may be close to $F_{0.1}$ (not estimated in the report). Combined with the literature and the bluefish SPR and YPR analyses, the review panel agrees that $SPR_{35\%}$ is a reasonable proxy of SSB_{MSY} . In my opinion this is consistent with the general life history of bluefish and past practices in management and in the history of exploitation.

Generic Research Track TOR 6. Define appropriate methods for producing projections; provide justification for assumptions of fishery selectivity, weights at age, maturity, and recruitment; and comment on the reliability of resulting projections considering the effects of uncertainty and sensitivity to projection assumptions.

This TOR had been met.

The WG conducted short-term projections under three F levels in WHAM, which incorporated auto-regressive processes and uncertainty in recruitment and numbers-at-age. The projections used the entire time series of recruitment (1985-2021), 5-year averages for natural mortality (assumed age varying but constant across years), maturity (constant), fishery selectivity, and weights-at-age. The life history study from TOR1 found that the maturity changes over time are limited, and the changes in weight-at-length were inconsequential.

The projection algorithm in dealing with the multi-fleet fishery matched the operational model setup. The review panel found the approach reasonable for projections under alternative fishing mortality level based on the council's risk policy and appropriate for management advice identification.

Generic Research Track TOR 7. Review, evaluate, and report on the status of research recommendations from the last assessment peer review, including recommendations provided by the prior assessment working group, peer review panel, and SSC. Identify new recommendations for future research, data collection, and assessment methodology. If any ecosystem influences from TOR 2 could not be considered quantitatively under that or other TORs, describe next steps for development, testing, and review of quantitative relationships and how they could best inform assessments. Prioritize research recommendations.

This TOR had been met.

The WG made considerable progress on several research recommendations stemming from the 2015 assessment (SAW/SARC 60) and MAFMC SSC recommendations in 2015 and 2021. Considerable progress in addressing these research recommendations was made and many of them were incorporated into the research track assessment model. The WG also developed four new research recommendations:

- Expand the collection of recreational release length frequency data. The recreational release length frequency should be spatially stratified; borrow if $n < 30$.
- Continue coastwide collection of length and age samples from fishery dependent and fishery independent sources.
- Continued development and refinement of the forage fish index; incorporate into the base model for management.
- Initiate fishery-independent or fishery dependent sampling programs to provide information on larger, older bluefish.

I agree with the above new research recommendations and those suggested by the panel at large. But of these, I am especially supportive of the last four where WHAM variability (process error) defines the scope of environmental/climate influences.

- Continue exploring the appropriate application of the WHAM model, including alternative ALK estimation.
- Explore the reasons for bimodal length frequency observed in bluefish harvest.
- Continue the forage fish index study and explore the potential application in catch rate standardization to remove the forage fish influence on catchability.
- Explore WHAM process error in simulating key parameter changes caused by climate or environmental changes, such as M and fish spatial distribution changes over time.

Generic Research Track TOR 8. Develop a backup assessment approach to providing scientific advice to managers if the proposed assessment approach does not pass peer review or the approved approach is rejected in a future management track assessment.

This TOR had been met.

The logic for the selection of the Ismooth (Legault et al. 2022) was clear. The BWG indicated this approach worked as well as any other Index Based method. The BWG also addressed reasons for not selecting other candidate approaches, including swept area, Depletion-corrected Average Catch (MacCall 2009), and Depletion-based Stock Reduction Analysis (Dick and MacCall 2011).

The Ismooth method is appropriate as a backup method even though it is not needed in this case. I recommend the WHAM model run for management purposes as the assessment goes forward.

Generic Research Track TOR 9. *Identify and consider any additional stock specific analyses or investigations that are critical for this assessment and warrant peer review, and develop additional TOR(s)* to address as needed.*

Note that this TOR was not within the WG's mandate and was not addressed explicitly by the WG or in the review meeting. However, it is still within the CIE's PWS for this review. Thus, I have used this TOR 9 as an opportunity to **identify and consider** possible avenues for more generic assessment research that would assist not only future assessments for this stock, but other stocks as well.

The WHAM framework provides an opportunity to explore ecological relationships with the environment and climate deviations. For example, a typical examination of some environmental index on recruitment is to assume that the index affects $R = Q \cdot aS / (1 + bS)$ or $R = Q \cdot aS \cdot \exp(-bS)$ where Q is a lognormal effect of the index on the density-independent factor (a). However, the population response is quite different when the Q is applied to the density-dependent factor (b). Similarly, the same sort of argument can be used with process error on M or NAA . If you are looking for population responses to climate/environment, perhaps the first step is to look to how a stock responds ecologically. WHAM gives one the flexibility to address that. Food for thought.

Conclusions and Recommendations

Spiny Dogfish

The spiny dogfish assessment moved from a "stochastic estimator" model to an SS3 implementation of sex-specific length-based dynamics. A thorough evaluation of catch, indices, life history growth and aging was conducted.

General lack of understanding of aging remains a weakness in the assessment. The SS3 configuration attempts to ameliorate the impacts somewhat, but it remains a problem. The model results show there are still inconsistencies. Nevertheless, it is recommended that the SS3 configuration be carried forward for management track evaluations. But it is recognized that some inconsistencies in SR model parameters, data weighting and uncertainties in catch will have to be reevaluated with the most recent data in the management track.

Bluefish

The bluefish research track assessment WG implemented a thorough examination of historical catches, indices, age-length keys, life history parameters. In so doing, it moved from the ASAP model to WHAM in a stepwise mode that was easy to follow and to justify.

There were a number of changes in data (most notably MRIPS recalibration), new and reevaluation of indices and updated natural mortality estimates, and the addition of selectivity blocks. The resulting WHAM model demonstrated minimal retrospective patterns in either F or

SSB. Results of the base model are similar to the previous assessment but with some shifts in scale.

The WHAM model as configured should provide a basis for management track evaluations going forward and is, thus, recommended.

Appendix 1: Bibliography of materials provided for review

Spiny Dogfish

Assessment Report • “Spiny_Dogfish_SAW_SARC_2022_FINAL.pdf”

Working Papers •

“Anstead 2022a.pdf” - Anstead K. 2022a. Natural mortality estimates for spiny dogfish.

“Anstead 2022b.pdf” - Anstead K. 2022b. Two data poor methods applied to spiny dogfish.

“Chang et al. 2022.pdf” - Chang J-H, Hart D and McManus MC. 2022. Stock synthesis for Atlantic spiny dogfish. •

“Hansell and McManus 2022.pdf” - Hansell A and McManus C. 2022. Spatio-temporal dynamics of spiny dogfish (*Squalus acanthias*) in US waters of the northwest Atlantic. •

“Hart and Chang 2022.pdf” - Hart DR, and Chang J-H. 2022. Per recruit modeling and reference points for spiny dogfish. •

“Hart and Sosebee 2022.pdf” - Hart DR, and Sosebee K. 2022. Length/Weight/Fecundity relationships for Atlantic spiny dogfish. •

“Jones 2022.pdf” - Jones AW. 2022. Exploring vessel trip report and observer based fishery information for spiny dogfish. •

“Jones et al. 2022.pdf” - Jones AW, Didden JT, McManus MC, and Mercer AJ. 2022. Exploring commercial CPUE indices for the spiny dogfish in the northeast U.S. •

“McCandless 2022.pdf” - McCandless C. 2022. Preliminary spiny dogfish movements and growth estimates from NEFSC mark recapture data. •

“McManus et al. 2022.pdf” - McManus MC, Sosebee K, and Rago P. 2022. Biological Reference Points for Spiny Dogfish: Revisiting Rago and Sosebee (2010). •

“Neiland and McElroy 2022.pdf” - Neiland JL and McElroy WD. 2022. NEFSC Gulf of Maine Bottom Longline Survey Data and Analyses for Spiny Dogfish •

“Passerotti and McCandless 2022.pdf” - Passerotti MS, and McCandless CT. 2022. Updated age and growth estimates for spiny dogfish *Squalus acanthias*. •

“Sosebee 2022a.pdf” - Sosebee KA. 2022a. Maturity of spiny dogfish in US waters from 1998- 2021. •

“Sosebee 2022b.pdf” - Sosebee KA. 2022b. Spiny dogfish catch summary and derivation of catch at length and sex.

Bluefish

The assessment document (Bluefish_SAW_SARC_2022_FINAL.pdf) includes all information to address the Terms of Reference, but we include additional detail in a series of working papers itemized below.

The working papers are available on the NEFSC data portal and on our github site.

“Assessment Report”

- Bluefish_SAW_SARC_2022_FINAL.pdf = Main assessment document

“Background”

- 01 readme.docx – document guide and repository of any report revisions.
- Background_documents.zip -> a small collection of grey and peer reviewed literature that the working group thought was especially helpful to accompany the assessment. Includes:
 - o The 2015 bluefish peer review report
 - o Stock and Miller 2021 -> describes the Woods Hole Assessment Model (WHAM) framework and software package
 - o Technical documentation for ASAP 3 -> technical documentation related to ASAP software used as part of early model exploration
 - o User manual for ASAP 3 -> technical documentation related to ASAP software used as part of early model exploration
 - o Ng et al. 2021 -> used as a guide for early development of the forage fish index
 - o Thorson 2019 -> used to help guide VAST modelling
- WP 01 Tyrell et al. 2022. Bluefish Ecosystem and Socioeconomic Profile.
- WP 02 Valenti 2022a. The Spatial Distribution of Bluefish (*Pomatomus saltatrix*): Insights from American Littoral Society Fish Tagging Data
- WP 03 Tyrell 2022. Bluefish VAST Index Exploration.
- WP 04 Gaichas et al. 2022. Vector Autoregressive Spatio-Temporal (VAST) modeling of piscivore stomach contents, 1985-2021.
- WP 05 Truesdell et al. 2022. Life History Analyses for Bluefish.
- WP 06 Tyrell and Truesdell 2022. Natural mortality of bluefish.
- WP 07 Celestino et al. 2022a. Index of abundance exploration and development by the Bluefish Working Group’s Fishery Independent Data Group.
- WP 08 Wood 2022a. TOR 2: Commercial and Recreational Data Collection and Analysis.
- WP 09 Drew 2022a. Recreational Data Changes for Bluefish, 2012-2021.
- WP 10 Drew 2022b. The Spatial Distribution of Bluefish (*Pomatomus saltatrix*): Insights from MRIP Data.
- WP 11 Valenti 2022b. Catch-and-Release Recreational Angling Mortality of Bluefish (*Pomatomus saltatrix*): Updated Analysis for 2022
- WP 12 Drew 2022c. Development of the Composite YOY Index for Bluefish.
- WP 13 Drew 2022d. A Fishery-dependent CPUE index for bluefish derived from MRIP data.
- WP 14 Celestino et al. 2022b. Development of Bluefish Age-Length Keys.
- WP 15 Wood 2022b. Bluefish Model Bridge-Building in ASAP.
- WP 16 Wood 2022c. ASAP diagnostic plots.
- WP 17 Wood 2022d. WHAM diagnostic plots.
- WP 18 Truesdell 2022. Alternative assessment plan.

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**

***Bluefish and Spiny Dogfish Research Track Peer Review
December 5-9, 2022***

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 standards¹.

Scope

The Research Track Peer Review 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 research track peer review is the cornerstone of the Northeast Region Coordinating Council stock assessment process, which includes assessment development, and report preparation (which is done by Working Groups or Atlantic States Marine Fisheries Commission (ASMFC) technical committees), assessment peer review (by the peer review panel), public presentations, and document publication. The results of this peer review will be incorporated

¹ https://www.whitehouse.gov/wp-content/uploads/legacy_drupal_files/omb/memoranda/2005/m05-03.pdf

into future management track assessments, which serve as the basis for developing fishery management recommendations.

The purpose of this meeting will be to provide an external peer review of the spiny dogfish and bluefish stocks. The requirements for the peer review follow. This Performance Work Statement (PWS) also includes: **Annex 1**: TORs for the research track, which are the responsibility of the analysts; **Annex 2**: a draft meeting agenda; **Annex 3**: Individual Independent Review Report Requirements; and **Annex 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. Modifications to the PWS and ToRs cannot be made during the peer review, and any PWS or ToRs modifications prior to the peer review shall be approved by the Contracting Officer's Representative (COR) and the CIE contractor. All TORs must be addressed in each reviewer's report. The reviewers shall have working knowledge and recent experience in the use and application of index-based, age-based, and state-space stock assessment models, including familiarity with retrospective patterns and how catch advice is provided from stock assessment models. In addition, knowledge and experience with simulation analyses and elasmobranchs is required.

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 Peer Review Panel (co)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 peer review meeting, using the criteria specified below in the “Tasks for Peer Review 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 peer review 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 Peer Review Panel 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 Peer Review Panel chair and reviewers combined:

Review the Report of spiny dogfish and bluefish Research Track Working Groups.

The Peer Review Panel 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 peer review meeting. For terms where a similar view can be reached, the Peer Reviewer Summary Report will contain a summary of such 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. Again, the CIE reviewers are not required to reach a consensus. The chair will take the lead in editing and completing this report. The chair may express their 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.

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 hybrid, both at the contractor’s facilities, and at the Northeast Fisheries Science Center in Woods Hole, Massachusetts, and via WebEx video conferencing. **CIE reviewers may attend virtually dependent on conditions of the COVID 19 pandemic**

Period of Performance

The period of performance shall be from the time of award through February, 2023. 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.

Within 2 weeks of award	Contractor selects and confirms reviewers
Approximately 2 weeks later	Contractor provides the pre-review documents to the reviewers
December 5-9, 2022	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 \$15,000.00.

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 Assessment Process Lead
 Northeast Fisheries Science Center
 166 Water Street, Woods Hole, MA 02543
Michele.Traver@noaa.gov

Annex 1. Generic Research Track Terms of Reference

1. Identify relevant ecosystem and climate influences on the stock. Characterize the uncertainty in the relevant sources of data and their link to stock dynamics. Consider findings, as appropriate, in addressing other TORs. Report how the findings were considered under the TORs.
2. Estimate catch from all sources including landings and discards. Describe the spatial and temporal distribution of landings, discards, and fishing effort. Characterize the uncertainty in these sources of data.
3. Present the survey data used in the assessment (e.g., indices of relative or absolute abundance, recruitment, state surveys, age-length data, application of catchability and calibration studies, etc.) and provide a rationale for which data are used. Describe the spatial and temporal distribution of the data. Characterize the uncertainty in these sources of data.
4. Use appropriate assessment approach to estimate annual fishing mortality, recruitment and stock biomass (both total and spawning stock) for the time series, and estimate their uncertainty. Compare the time series of these estimates with those from the previously accepted assessment(s). Evaluate a suite of model fit diagnostics (e.g., residual patterns, sensitivity analyses, retrospective patterns), and (a) comment on likely causes of problematic issues, and (b), if possible and appropriate, account for those issues when providing scientific advice and evaluate the consequences of any correction(s) applied.
5. Update or redefine status determination criteria (SDC; point estimates or proxies for BMSY, BTHRESHOLD, FMSY and MSY reference points) and provide estimates of those criteria and their uncertainty, along with a description of the sources of uncertainty. If analytic model-based estimates are unavailable, consider recommending alternative measurable proxies for reference points. Compare estimates of current stock size and fishing mortality to existing, and any redefined, SDCs.
6. Define appropriate methods for producing projections; provide justification for assumptions of fishery selectivity, weights at age, maturity, and recruitment; and comment on the reliability of resulting projections considering the effects of uncertainty and sensitivity to projection assumptions.
7. Review, evaluate, and report on the status of research recommendations from the last assessment peer review, including recommendations provided by the prior assessment working group, peer review panel, and SSC. Identify new recommendations for future research, data collection, and assessment methodology. If any ecosystem influences from TOR 2 could not be considered quantitatively under that or other TORs, describe next steps for development, testing, and review of quantitative relationships and how they could best inform assessments. Prioritize research recommendations.
8. Develop a backup assessment approach to providing scientific advice to managers if the proposed assessment approach does not pass peer review or the approved approach is rejected in a future management track assessment.

9. Identify and consider any additional stock specific analyses or investigations that are critical for this assessment and warrant peer review, and develop additional TOR(s)* to address as needed.

Research Track TORs:

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

Guidance to Peer Review Panels 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, $OFL \geq ABC$.]

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 peer review 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.

Annex 2. Draft Review Meeting Agenda

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

Spiny Dogfish and Bluefish Track Assessment Peer Review Meeting

December 5-9, 2022

WebEx link: TBD

DRAFT AGENDA* (v. 6/21/2022)

**All times are approximate, and may be changed at the discretion of the Peer Review Panel 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 Peer Review Panel.*

Monday, December 5, 2022

Time	Topic	Presenter(s)	Notes
9 a.m. - 9:30 a.m.	Welcome/Logistics Introductions/Agenda/ Conduct of Meeting	Michele Traver, Assessment Process Lead Russ Brown, PopDy Branch Chief Panel Chair	
9:30 a.m. - 10:30 a.m.	TOR #1		Spiny dogfish
10:30 a.m. - 10:45 a.m.	Break		
10:45 a.m. - 11:45 a.m.	TOR #2		
11:45 a.m. - 12:15 p.m.	Discussion/Summary	Review Panel	
12:15 p.m. - 12:30 p.m.	Public Comment	Public	
12:30 p.m. - 1:30 p.m.	Lunch		
1:30 p.m. - 3 p.m.	TOR #3		
3 p.m. - 3:15 p.m.	Break		
3:15 p.m. - 4:15 p.m.	TOR #4		
4:15 p.m. - 4:45 p.m.	Discussion/Summary	Review Panel	
4:45 p.m. - 5 p.m.	Public Comment	Public	
5 p.m.	Adjourn		

Tuesday, December 6, 2022

Time	Topic	Presenter(s)	Notes
9 a.m. - 9:15 a.m.	Welcome/Logistics	Michele Traver, Assessment Process Lead Panel Chair	Spiny dogfish cont.
9:15 a.m. - 10:30 a.m.	TOR #5		
10:30 a.m. - 10:45 a.m.	Break		
10:45 a.m. - 11:45 a.m.	TOR #6		
11:45 a.m. - 12:15 p.m.	Discussion/Summary	Review Panel	
12:15 p.m. - 12:30 p.m.	Public Comment	Public	
12:30 p.m. - 1:30 p.m.	Lunch		
1:30 p.m. - 3 p.m.	TOR #7		
3 p.m. - 3:15 p.m.	Break		
3:15 p.m. - 4:15 p.m.	TOR #8-9		
4:15 p.m. - 4:45 p.m.	Discussion/Summary	Review Panel	
4:45 p.m. - 5 p.m.	Public Comment	Public	
5 p.m.	Adjourn		

Wednesday, December 7, 2022

Time	Topic	Presenter(s)	Notes
9 a.m. - 9:15 a.m.	Welcome/Logistics	Michele Traver, Assessment Process Lead Panel Chair	Bluefish
9:15 a.m. - 10:30 a.m.	TOR #1		
10:30 a.m. - 10:45 a.m.	Break		
10:45 a.m. - 11:45 a.m.	TOR #2		
11:45 a.m. - 12:15 p.m.	Discussion/Summary	Review Panel	
12:15 p.m. - 12:30 p.m.	Public Comment	Public	

12:30 p.m. - 1:30 p.m.	Lunch		
1:30 p.m. - 3 p.m.	TOR #3		
3 p.m. - 3:15 p.m.	Break		
3:15 p.m. - 4:15 p.m.	TOR #4		
4:15 p.m. - 4:45 p.m.	Discussion/Summary	Review Panel	
4:45 p.m. - 5 p.m.	Public Comment	Public	
5 p.m.	Adjourn		

Thursday, December 8, 2022

Time	Topic	Presenter(s)	Notes
9 a.m. - 9:15 a.m.	Welcome/Logistics	Michele Traver, Assessment Process Lead Panel Chair	Bluefish
9:15 a.m. - 10:30 a.m.	TOR #5		
10:30 a.m. - 10:45 a.m.	Break		
10:45 a.m. - 11:45 a.m.	TOR #6		
11:45 a.m. - 12:15 p.m.	Discussion/Summary	Review Panel	
12:15 p.m. - 12:30 p.m.	Public Comment	Public	
12:30 p.m. - 1:30 p.m.	Lunch		
1:30 p.m. - 3 p.m.	TOR #7		
3 p.m. - 3:15 p.m.	Break		
3:15 p.m. - 4:15 p.m.	TOR #8-9		
4:15 p.m. - 4:45 p.m.	Discussion/Summary	Review Panel	
4:45 p.m. - 5 p.m.	Public Comment	Public	
5 p.m.	Adjourn		

Friday, December 9, 2022

Time	Topic	Presenter(s)	Notes
-------------	--------------	---------------------	--------------

9 a.m. - 5 p.m.	Report Writing	Review Panel	
-----------------	----------------	--------------	--

Annex 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.

Annex 4. Peer Reviewer Summary Report Requirements

1. The main body of the report shall consist of an introduction prepared by the Research Track Peer Review Panel 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 peer review meeting. 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. It should also include whether they accept or reject the work that they reviewed, with an explanation of their decision (strengths, weaknesses of the analyses, etc.)

To make this determination, the peer review panel chair and reviewers should consider whether or not the work provides a scientifically credible basis for developing fishery management advice. If the reviewers and peer review panel 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 peer review meeting, and relevant papers cited in the Peer Reviewer Summary Report, along with a copy of the CIE Performance Work Statement.
4. The report shall also include as a separate appendix the assessment Terms of Reference used for the peer review meeting, including any changes to the Terms of Reference or specific topics/issues directly related to the assessments and requiring Panel advice.

Appendix 3: Panel membership or other pertinent information from the panel review meeting.

**Spiny Dogfish/Bluefish Research Track Peer Review Attendance
December 5-9, 2022**

GARFO - Greater Atlantic Regional Fisheries Office
MADMF - Massachusetts Division of Marine Fisheries
MAFMC - Mid Atlantic Fisheries Management Council
MDNR - Maryland Department of Natural Resources
NEFMC - New England Fisheries Management Council
NEFSC - Northeast Fisheries Science Center
NCDMF - North Carolina Division of Marine Fisheries
NJFW - New Jersey Fish and Wildlife
NYSDEC - New York State Department of Environmental Conservation
RIDEM - Rhode Island Department of Environmental Management

~~~~~  
*Yan Jiao - Chair*

*Joe Powers - CIE Panel*

*Robin Cook - CIE Panel*

*Paul Medley - CIE Panel*

Russ Brown - NEFSC, Population Dynamics Branch Chief

Michele Traver - NEFSC, Assessment Process Lead

Abby Tyrell - NEFSC

Alan Bianchi - NCDMF

Alex Dunn - NEFSC

Alex Hansell - NEFSC

Alexei Sharov - MDNR

Andy Jones - NEFSC

Anna Mercer - NEFSC

Brandon Muffley - MAFMC staff

Brian Linton - NEFSC

Cami McCandless - NEFSC

Charles Adams - NEFSC

Charles Perretti - NEFSC

Chris Legault - NEFSC

Conor McManus - RIDEM

Cynthia Ferrio - GARFO

Dave McElroy - NEFSC

Dvora Hart - NEFSC

Eric Robillard - NEFSC

Greg DiDomenico - Lunn's Fisheries

Hannah Hart - MAFMC staff



James Fletcher - United National Fishermen's Association  
Jason Didden - MAFMC staff  
John Maniscalco - NYSDEC  
Jose Montanez - MAFMC staff  
Jui-Han Chang - NEFSC  
Julie Nieland - NEFSC  
Karson Cisneros - MAFMC staff  
Kathy Sosebee - NEFSC  
Katie Drew - ASMFC staff  
Kiersten Curti - NEFSC  
Kristen Anstead - ASMFC  
Larry Alade - NEFSC  
Liz Brooks - NEFSC  
Mark Terceiro - NEFSC  
Mike Celestino - NJFW  
Michelle Passerotti - NEFSC  
Paul Nitschke - NEFSC  
Rich McBride - NEFSC  
Ricky Tabandera - NEFSC  
Sam Truesdell - MADMF  
Samantha Werner - NEFSC  
Sarah Gaichas - NEFSC  
Scott Large - NEFSC  
Tim Miller - NEFSC  
Toni Chute - NEFSC  
Tony Wood - NEFSC