11.D.1(2)
182nd CM

WESTERN PACIFIC STOCK ASSESSMENT REVIEW

"Stock assessment of uku (*Aprion virescens*) in Hawaii, 2020"

Individual Reviewer Report

by
Erik C. Franklin (Chair)

March 25, 2020

Prepared for
Pacific Islands Fisheries Science Center, National Marine Fisheries Service, NOAA
Pacific Islands Regional Office, National Marine Fisheries Service, NOAA
Western Pacific Regional Fishery Management Council

Table of Contents

Executive Summary

A Western Pacific Stock Assessment Review (WPSAR) of the 2020 Main Hawaiian Islands uku (*Aprion virescens*) benchmark stock assessment was performed in Honolulu, HI during February 24-28, 2020. The purpose of the WPSAR was to evaluate the utility and applicability of the data, CPUE standardization, assessment models, model parameters, uncertainty, and assumptions presented in the assessment report for making recommendations on the status determination criteria of uku. The WPSAR would also provide recommendations to improve the current stock assessment and future work. This document represents the WPSAR Chair's individual review report.

The content of this review addresses the Terms of Reference applied to the integrated statistical catch-at-age stock assessment of uku *Aprion virescens* in the main Hawaiian Islands (MHI). The uku assessment incorporated population abundance indices (i.e., CPUE), catch, and size composition data from fishery-dependent catch statistics and fishery-independent diver surveys with life history parameters for longevity, growth, survivorship, and maturity from local studies. The foundation of the assessment approach was the use of the integrated statistical catch-at-age assessment model (i.e., Stock Synthesis v 3.30) that generates maximum likelihood estimates of population parameters, outputs, and uncertainty. The assessment utilized commercial data from fisher reports that covered the period 1948-2018 for deep-sea handline (dominant gear) and trolling and inshore handline gear for 2003-2018. The recreational data from HMRFSS covered the period 2003-2018 with data reconstructions for 1948-2002. Fishery-independent diver surveys were performed during six years (i.e., 2005, 2008, 2010, 2012, 2015, 2016). Life history parameters for length-at-age, longevity, and maturity of uku were determined from studies on the Hawaiian uku population. All materials for the review were provided by the WPSAR coordinating committee in a timely fashion. The assessment team deserves praise for the effort that went into the assessment report, their timely responses to requests for additional analyses and extra information during the review, and their collegial attitude and interactions during the entire review process.

The base case assessment presented the status of uku in the Main Hawaiian Islands (MHI) as not overfished, not experiencing overfishing. Based on the results contained in the assessment report and additional sensitivity runs presented during the review panel, I recommend that the base case assessment is based on the best scientific information available and can be used to make a status determination for the uku stock in the MHI. Over twenty alternative model scenarios were presented to evaluate the sensitivity of results to variability in life history parameters, catch data, CPUE indices, and size composition data which demonstrated a remarkable stability in model outcomes relative to reference points. I found that the assessment report sufficiently met the Terms of Reference (TOR) and provide a detailed response to describe how each TOR was satisfied. To improve the final version of the assessment report and guide future scientific activities, I include a list of short, medium, and long-term recommendations. I envision that the stock assessment strategy and framework utilized for MHI uku can serve as a template for future single species stock assessments in the Western Pacific region.

Background

The green jobfish *Aprion virescens* known in Hawaii as "uku" is a bottomfish management unit species (BMUS) federally managed under the "Fishery Ecosystem Plan for the Hawaii Archipelago" by the Western Pacific Regional Fishery Management Council (WPRFMC). The uku has a long history of fishing exploitation in the Hawaiian Islands but was historically assessed as not being overfished (Sabater and Kleiber 2013) and not experiencing overfishing (Nadon 2017). A more recent benchmark stock assessment for uku was performed to update the status of the fishery by Nadon et al. (2020). This document provides an individual reviewer report of the Western Pacific Stock Assessment Review (WPSAR) of the "Stock assessment of uku (*Aprion virescens*) in Hawaii, 2020" by M. Nadon, M. Sculley, and F. Carvalho. This WPSAR addresses a set of eleven (11) Terms of Reference (TOR) for the review of a benchmark stock assessment of the green jobfish or uku (*Aprion virescens* Valenciennes, 1830) in the main Hawaiian Islands, following guidelines established in the WPSAR framework. The WPSAR framework identifies a peer review process for the scientific information used to advise the WPRFMC about the conservation and management of their fisheries. Brief descriptions of the TORs are: (1) TOR 1 is the appropriateness of the data used in the assessment; (2) TOR 2 is the appropriateness of the CPUE standardization; (3) TOR 3 is the appropriateness of the assessment models; (4) TOR 4 relates to decision points and input parameters; (5) TOR 5 relates to sources of uncertainty; (6) TOR 6 relates to model assumptions; (7) TOR 7 relates to scientific validity of results; (8) TOR 8 relates to the adequacy of the projection methods; (9) TOR 9 relates to the validity of the results to address management goals; (10) TOR 10 addresses recommendations for improvements and future research, and (11) TOR 11 is a request for the individual panelist reports and chair summary report. A list of the fully detailed TORs is provided in the Appendix.

This report includes a description of the role and tasks of the WPSAR Chair, general comments on the assessment, detailed responses to each of the Terms of Reference, and a synopsis on the consensus summary findings of the panel.

WPSAR Chair Role and Tasks

For the MHI uku assessment WPSAR panel, I served in the role of WPSAR Chair and completed the following tasks under sections *Pre-review Background Documents*, *Panel Review Meeting*, and *Post Panel Review Tasks* in accordance with the Statement of Work provided by the WPSAR Coordinating Committee.

Pre-review Background Documents

While serving as Chair for this WPSAR panel, I read all the following documents prior to the in-person peer review during February 24-28, 2020. Required pre-review documents:

- DRAFT 2020 uku assessment: Nadon et al. Stock assessment of uku in Hawaii, 2020. NOAA Tech Memo.

- Previous reef fish stock assessment: Nadon, M. O. 2017. Stock assessment of the coral reef fishes of Hawaii. U.S. Dep. Commerce, NOAA Tech. Memo., NOAA-TM-NMFS-PIFSC-60, 212 p. (focus on sections pertaining to uku)
- Independent peer review report for Nadon (2017) stock assessment: Benchmark review of the 2016 stock assessment of the Main Hawaiian Islands Reef-associated Fish. Consensus Review Panel Report. 27 p.
- Hawaii Fishery Ecosystem Plan: Western Pacific Regional Fishery Management Council. 2009. Fishery Ecosystem Plan for the Hawaii Archipelago. Only section 5.2 (pp. 133-138) and section 5.3 (pp. 138-143).
- Methot, R.D. and Wetzel, C. 2013. Stock synthesis: A biological and statistical framework for fish stock assessment and fishery management. Fisheries Research 142: 86-99.
- Winker et al. 2014. Proof of concept for a novel procedure to standardize multispecies catch and effort data. Fisheries Research 155: 149-159.

Panel Review Meeting

The panel review meeting was held February 24, 2020 through February 28, 2020 in the Western Pacific Regional Fishery Management Council Office, Suite 1401, Finance Factors Building, 1164 Bishop Street, Honolulu, Hawaii 96813. An agenda of the review meeting and list of participants, presenters, and attendees are in the Appendix. During the panel review meeting, I conducted an independent peer review, along with two Center for Independent Experts (CIE) reviewers, Dr. Yong Chen (University of Maine) and Dr. Yan Jiao (Virginia Tech), in accordance with the Statement of Work (SOW) and Terms of Reference (TORs). As Chair, I facilitated the review to accomplish the stated goals and objectives articulated in the WPSAR SOW and TORs. I led discussions with the assessment team in seeking clarification and requested additional analyses to address the TORs. At the conclusion of each review day, I conferred with members of the review panel to summarize findings, discussed comments, and attempted to arrive at a consensus relative to each of the TOR elements discussed that day. During the appropriate time, I sought and received public comments. At the conclusion of the review, I produced a draft Consensus Summary Report outlining the consensus recommendations from the three panel members addressing each TOR. I presented this report on the last day of the review (February 28, 2020) to the assessment team, WPSAR principals, and the public.

Post Panel Review Tasks

I completed an independent peer review report (i.e., this document) in accordance with the WPSAR SOW. I also completed a Consensus Summary Report with input from all panel members to describe findings for each TOR. I provided a brief synopsis of my views on the summary findings and conclusions reached by the panel (in this document). I plan to present, in-person, the consensus results of the review to the SSC and Council after finalization of the reviewed assessment document.

General Review Comments

The stock assessment authors are to be commended for a well written document that describes their methods of analysis, model results and uncertainty, and scientific recommendations for the management of uku *Aprion virescens* in the main Hawaiian Islands. This work represents a strong progression from the prior data-poor assessments (Sabater and Kleiber 2013, Nadon 2017) to an integrated statistical catch-at-age approach for uku (Nadon et al. 2020) that should provide a framework for the future assessment of other managed species in the US Western Pacific region. A list of general comments that I recorded during an initial review of the assessment follow.

- The title of the assessment is not sufficiently informative. I suggest the alternative title, ”Benchmark Stock Assessment of Uku (*Aprion virescens*) in the Main Hawaiian Islands, 2020” to better represent whether this is a benchmark or update assessment and more accurately describe the geographic extent of the exploited stock (i.e., it does not include the Northwestern Hawaiian Islands which is implied by using only “Hawaii” in the title).
- The total commercial catch was provided by the Division of Aquatic Resources but this is different than the catch provided by WestPacFin that compares the fisher reports with dealer reports and cleans the dataset to create a final.
- Size composition data was drawn from the deep-sea handline data set only. Others gears were reported as having insufficient reporting data to construct annual size frequencies but there is no presentation of these other data sets.
- Section 5.3.5 might be better titled “Assessment Strategy and Sensitivity Analyses” or a similar title as “Assessment Strategy” isn’t informative enough.
- There is a level of reassurance that three separate assessments (Sabater and Kleiber 2013, Nadon 2016, Nadon et al. 2020) have all concluded that the uku MHI stock is not overfished, nor experiencing overfishing. Taken with a grain of salt given the quality of catch data inputs for the stock but still reassuring.
- The recent positive trend in CPUE from 2002 to 2018 is interesting. It coincides with the shift in effort units in the indices time series. Is this a real trend reflecting an increase in stock biomass or is this a ‘hyperstable’ CPUE driven by targeting the Penguin Banks populations with ever more efficient gears?
- Fig 23, the selectivity for the diver surveys corresponds to the troll selectivity but on page 15 the selectivity is stated as coming from the “recreational sector”.
- The methods and data sources are reasonably detailed, but the work is not currently reproduceable as presented. I suggest that the authors include an appendix or a supplementary technical report that explicitly includes the data inputs, code for the analysis and visualization scripts, and sufficient guidelines to reproduce the results presented in the assessment.
- The github repository (https://github.com/PIFSCstockassessments/MHI-Uku-2020-Final) listed in the assessment is not publicly available.

Responses to TORs

Detailed responses to the Term of Reference are given in the following sections. A summary table provides an overview of the responses (Table 1).

Table 1. Summary of responses to Term of Reference 1-9 (TOR). Potential responses were "Yes", "Yes with caveats", or "No". TOR 10 were recommendations and TOR 11 is this report.

TOR	Response
1: Data	Yes, with caveats
2: CPUE Standardization	Yes
3: Models	Yes
4: Decision Points and Input Parameters	Yes
5: Uncertainty	Yes, with caveats
6: Assumptions	Yes
7: Scientific Validity	Yes
8: Projections	Yes, with caveats
9: Management	Yes

TOR 1 Data: *Of the data considered for inclusion in the assessment, were final decisions on inclusion/exclusion of particular data appropriate, justified, and well-documented?*

Yes with caveats, the final decisions on inclusion/exclusion of particular data were appropriate, justified, and well-documented.

In general, the data included in the assessment were appropriate, justified and well-documented. The assessment team compiled fishery-dependent datasets of commercial catch and effort from the state of Hawaii Division of Aquatic Resources (DAR) fisher reporting system (FRS) that provided CPUE indices for: (1) deep-sea handline (DSH) from 1948 to 2002 with effort reported as fishing days, as well as DSH, inshore handline (ISH) and trolling (TROL) from 2003 to 2018 with effort reported as hours fished. Non-commercial catch was used from the Hawaii Marine Recreational Fishing Survey (HMRFS) for data from 2003 to 2018 and a model-based reconstruction was used for data from 1948 to 2002 A fishery-independent diver survey was also used to provide an index of relative abundance for the assessment for a set of 6 years from 2005 to 2016. Life history parameters were generated from local studies of age and growth for the species. Explicit and appropriate criteria were developed to filter the data to generate final data sets for the assessment. While the assessment team did an excellent job in evaluating, preparing, and filtering these data for the assessment, there are a number of concerns that should be highlighted.

The commercial fisher reporting system relies on self-reported catch and effort data from fisherman. While the data may suffer from reporting inaccuracies due to self-reporting, the data collection protocols included quality control and evaluation systems (e.g., checking fisher reports with dealer reports). The DSH data also provides an enviable time series of 70 years but it was unclear how fishing power changed over the time period. Given these explicit shortcomings, the data should still be used for an assessment as it represents the primary set of information to evaluate the status of the stock. The ISH data included in the assessment was from 2003 to 2018 but it appears as if there may be additional years (i.e., 1990 to 2002) to include in the assessment (Fig. 1).The size composition data used in the assessment was drawn from the DSH gear only. We were also given background by an active fisherman on the evolving technology that can improve gear efficiency and fishing strategy among the commercial and non-commercial fleet. It was unclear if these new elements in the fishery were adequately addressed in the current assessment but should be considered for future data collection and assessments.

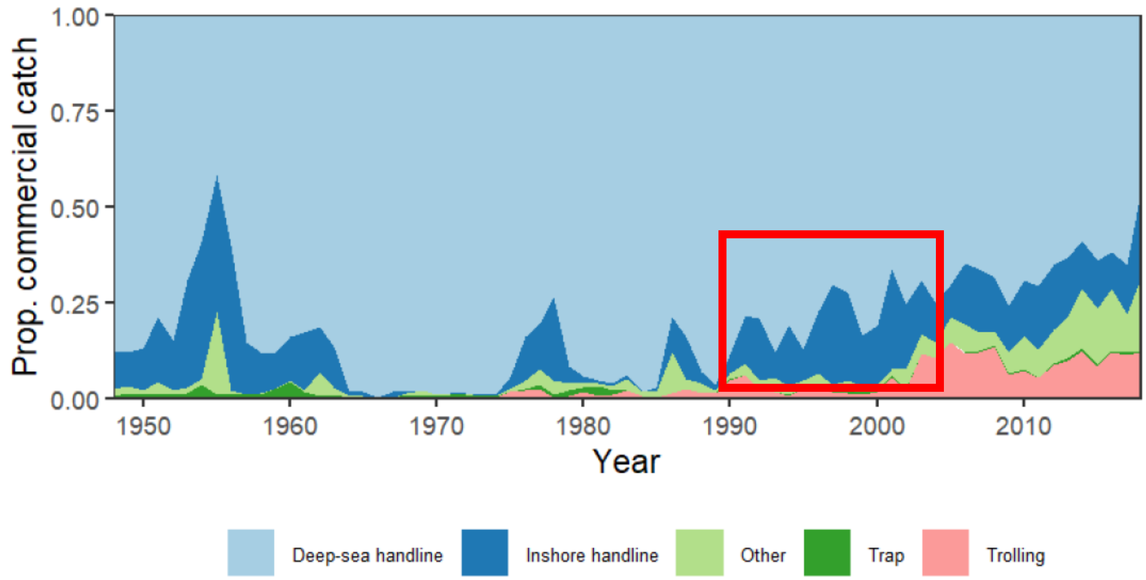

Fig 1. Proportion of total commercial catch caught by each gear time (Figure 4 in assessment). Inshore handline catch from 1990-2002 increases relative to prior years (red box) but these data were not included in the assessment.

The non-commercial (i.e., recreational) data presents multiple issues. First, the interannual variability in the HMRFS data (from 2003 to 2018) suggests needed improvement in the sampling design and data collection protocols (Fig. 2). Recent estimates (for 2013-2018) suggest that the recreational sector is responsible for 33-73% of total annual catch in the fishery (calculated from data in Table 2 of the assessment). If those estimates of recreational catch relative to the commercial catch are accurate then a concerted effort needs to be made to overhaul the HMRFS program to provide better annual estimates of non-commercial catch. Next, the historical reconstruction method used for recreational data for 1948-2002 is weak and unable to reflect any inaccuracies or uncertainty around the estimated catches given the assumed differences in the fishery between the reference period (2003-2007) and the reconstruction period (1948-2002). The approach assumes that the ratio of catch to human population is constant for the reference period and the reconstruction period. Alternate approaches were pursued but none were preferable. While these issues present some flaws in the recreational data, it appears to represent an important sector of the uku fishery that needs to be addressed in the assessment but should be a priority for significant improvements in sampling design and data collection.

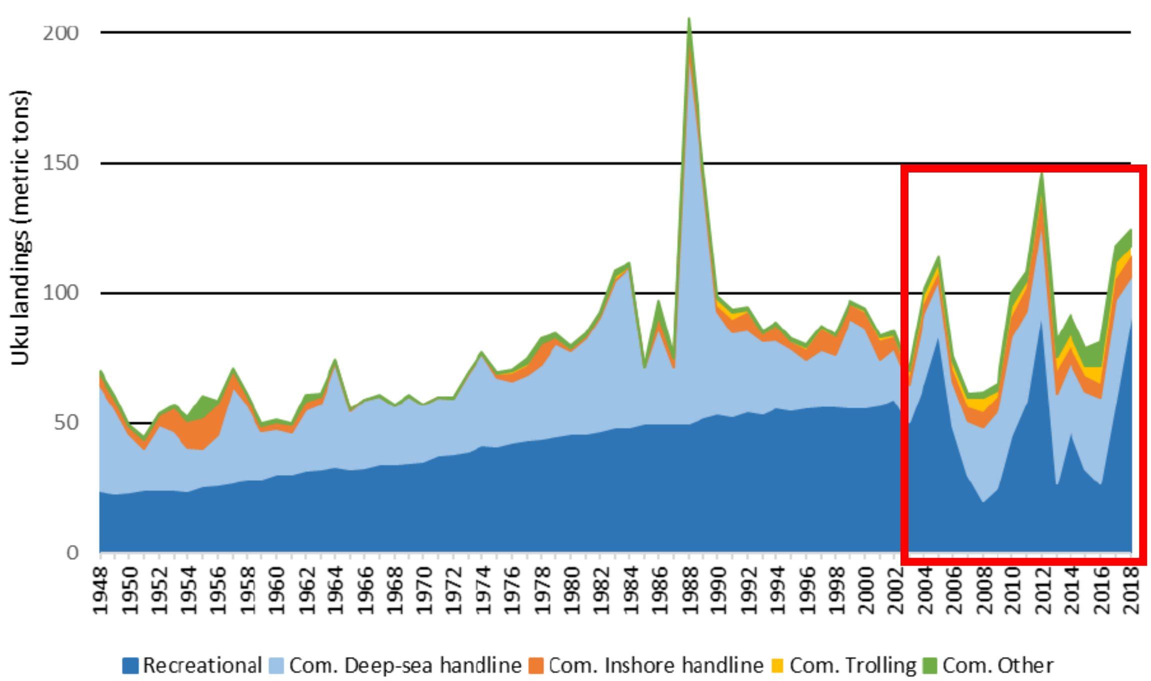

Figure 2. Historical uku catch by sector (Figure 6 in assessment). The period of extreme interannual variability in estimated recreational catch is data estimated from HMRFS. Prior years for the recreational data were generated by a historical reconstruction model.

The fishery-independent diver surveys provide a valuable potential index of abundance for the shallow (< 30 m depth) portion of the uku stock. A well-designed survey can provide an unbiased relative index of abundance for this species, but the diver surveys should be interpreted with some caution given the inconsistent geographic sampling and multispecies focus of the methods which introduces high uncertainty. The time series is also scant compared to the other datasets with six survey years across 2005 to 2016. Given the sparsity of data, the assessment results did not change appreciably between a sensitivity run without the diver survey performed during the review and the base case (Fig. 3). While these data have significant issues, annual surveys with a sampling design and methods focused on uku would greatly improve the quality of this data.

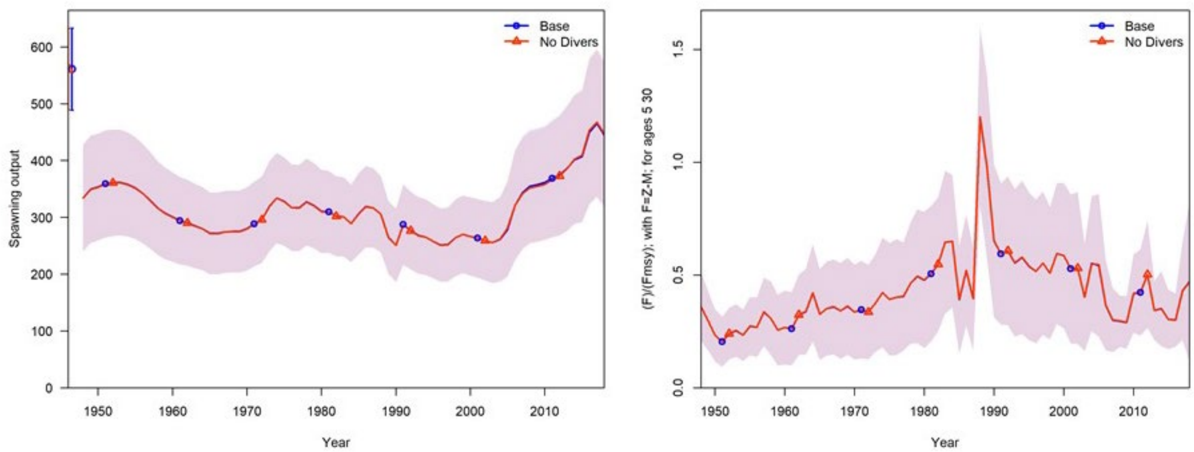

Figure 3. A comparison of model outputs (SSB and F/FMSY) for the base case and a sensitivity run that did not include the fishery-independent diver survey index.

The life history data used in the assessment were generated from a local, unpublished study performed by the NOAA PIFSC Life History Program. The processing and reading of the otoliths followed standard procedures for quality assurance (J. O'Malley pers comm). The definition of a stock here is a bit muddled as the assessment defines the stock boundaries as the Main Hawaiian Islands (Figure 3 in assessment) but use an estimate of longevity from the northwestern Hawaiian Islands (NWHI). There is some evidence to suggest that there is biological separation of some stocks between the MHI and NWHI (Toonen et al 2011), but there is not general acceptance of the concept in the literature as some demersal fish species that have no genetic population structure across the entire Hawaiian Archipelago (Craig et al. 2011). The growth models between the MHI and NWHI were not statistically different, as communicated by the assessment team (Fig. 4). One potential concern with the inclusion of the NWHI in the growth model, if there are two stocks, is that the longevity is different between the two regions (MHI = 27 years, NWHI = 32 years) which would affect the estimate of natural mortality. There may be a biological basis for the difference, or it could be due to a truncation of the MHI stock size structure by fishing.

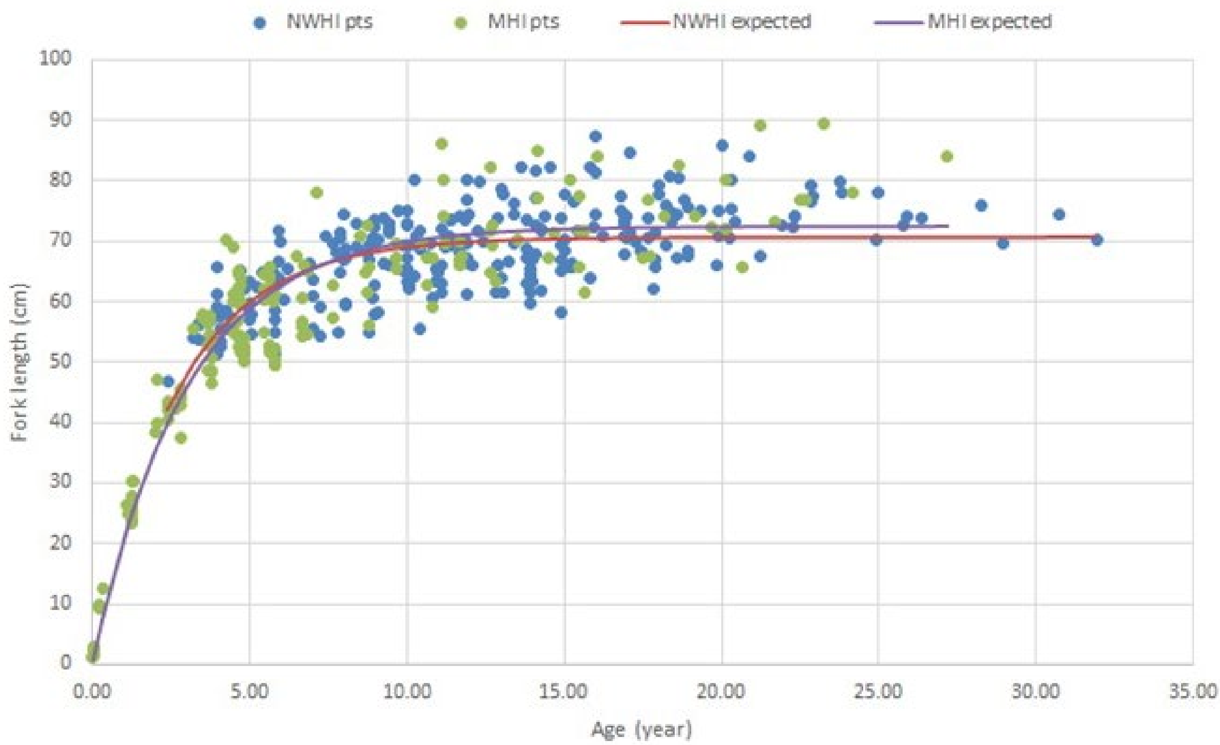

Figure 4. Uku age-at-length data from specimens processed by the NOAA PIFSC LHP. The lines are von Bertalanffy growth models fit to MHI and NWHI data. There was no reported statistical differences between models (Figure prepared by stock assessment team during review).

TOR 2 CPUE Standardization: *Is the CPUE standardization properly applied and appropriate for this species, fishery, and available data?*

Yes, the CPUE standardization was applied properly and appropriate for this species, fishery, and available data.

In general, the assessment provided a clear structure and a detailed sequence of actions used to standardize CPUE indices for uku-targeted records that was properly applied and appropriate. The use of fishery-dependent CPUE data for uku may not accurately reflect the status of stock abundance accuracy (Maunder et al. 2006), these data represent the best scientific information available to assess long-term trends in the dynamics of stock. Several issues related to data cleaning (i.e., tracking individual license holders through the time series, multi-day trips, selecting uku targeted trips, and identifying covariates affecting CPUE) were reconciled by using methods adopted from prior bottomfish data workshops (Yau 2018) and the most recent deep 7 bottomfish assessment (Langseth et al. 2018). Due to high proportion of zero catch records, delta lognormal GLMMs were used to standardize CPUE. Model variable selection and diagnostic evaluations were appropriate but the panel recommended exploring a model run with the interaction of month:PC since catch compositions were said to shift seasonally with uku being catch more frequently during summer months. Of greater concern was a coincidental shift in the direction of the trends in standardized CPUE from negative to positive that occurred when the unit of effort changes from "fishing days" to "fishing hours" in 2003 (Figure 28 in the assessment). To address this concern, assessment team performed a sensitivity run of the DSH CPUE using just "fishing days" for the 1948 to 2018 time series (Fig. 5) which demonstrated the same shifting trend suggesting that the trend was not an artifact of the change in unit of effort. The assessment team also compared the standardized CPUE indices using either effort as "fishing days" or "fishing hours" for three gears during the 2003-2018 period and found negligible differences in the values and trends between the two effort units (Fig. 6). These results suggest that the use of fishing hours is relatively comparable to fishing days for the CPUE time series.

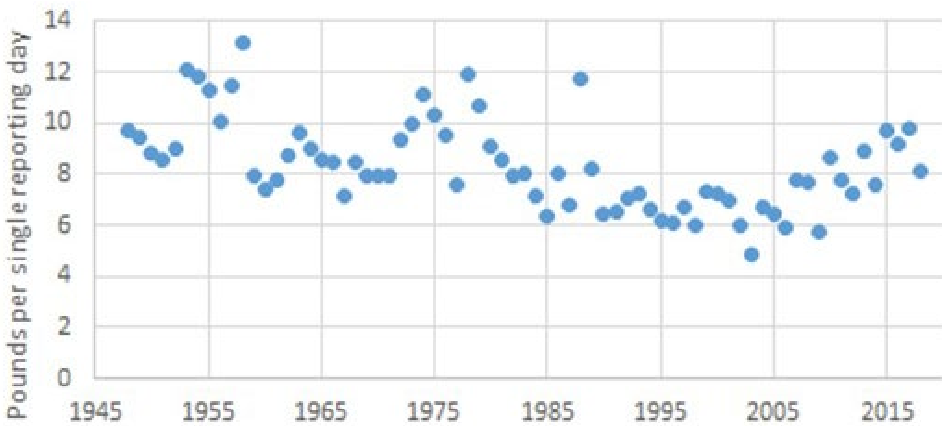

Figure 5. Standardized deep-sea handline CPUE calculated with "fishing days" as the unit of effort throughout the entire time series. In the assessment, the time series shifted to "fishing hours" as the unit of effort and demonstrated a similar trend in CPUE. Prepared by the assessment team.

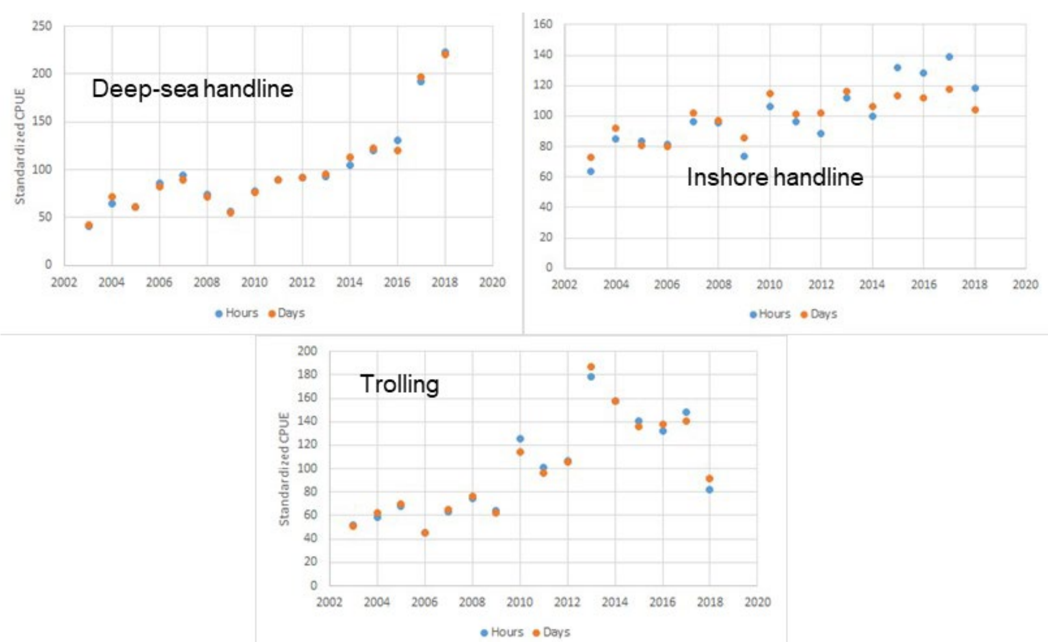

Figure 6. A comparison of standardized CPUE indices for three gears calculated with "fishing days" or "fishing hours" as the units of effort. Prepared by the assessment team.

TOR 3 Models: *Are the assessment models used reliable, properly applied, adequate, and appropriate for the species, fishery, and available data?*

Yes, the assessment models used are reliable, properly applied, adequate, and appropriate for the species, fishery, and available data.

The primary model platform used for the stock assessment was Stock Synthesis v3.30 (SS3), a standard and reliable integrated statistical catch-at-age model. The assessment incorporated multiple CPUE indices, size composition data, and life history parameters to model the status of the uku stock in the main Hawaiian Islands (MHI). The SS3 configurations were properly applied and adequate to model the growth, maturation, steepness, natural mortality, size composition, selectivity, and catchability. The use of the LBSPR model approach (Hordyk et al. 2016) to generate selectivity parameters for inshore handline, trolling, and other gears from limited size composition data was a creative and acceptable solution for these data poor gears (e.g., Figs 22-23 in the assessment). An adequate and properly applied set of model diagnostics included an assessment of model convergence using several criteria (i.e., Hessian inversion, low max gradient component, jittering start values), analyses of residuals to CPUE indices, the construction of a likelihood profile for R_0, and a retrospective analysis. The series of visualizations for the model outputs and diagnostics were appropriate and adequate to document the results from the base case assessment and sensitivity runs.

TOR 4 Decision Points and Input Parameters: *Are decision points and input parameters reasonably chosen?*

Yes, decision points and input parameters were reasonably chosen.

The assessment (and the assessment team during the review) provided consistent and clear rationales for their choices on decision points and input parameters. For the data inputs, the assessment described the methods used to filter appropriate data and generate CPUE indices for the different gear types. The selectivities for all gears but the DSH (estimated in SS3) were determined using the LBSPR method which was a reasonable approach given the paucity of size composition data. The length-at-age data had a sufficient sample size to estimate growth parameters and relied on standard diagnostics to evaluate the aging criteria. Longevity from a local growth study was used to estimate natural mortality and apply it as a fixed parameter in the model. An alternative approach could be to apply an age-specific mortality rate (Lorenzen m) to the model. This suggestion from the panel was evaluated by the assessment team and showed no appreciable difference in selectivity for DSH using the fixed M or age-specific M (Lorenzen M in SS3). The framework from the assessment strategy and sensitivity runs included in the assessment provided a logical and transparent approach to determine the components that comprised the base case assessment which represents the best scientific information available to assess the MHI uku stock.

TOR 5 Uncertainty: *Are primary sources of uncertainty documented and presented?*

Yes, with caveats, the primary sources of uncertainty were documented and presented.

In general, the sources of uncertainty for the data, model, and input values were adequately documented and described in the assessment. The log-likelihood profiles, CPUE residual analysis, and retrospective analysis results provided standard diagnostics to document the performance of the model and data. It's unclear if the recent positive trend in CPUE reflects a relative increase in stock biomass or ever more efficient gear from technology improvements focused on a summer spawning population based on Penguin Bank. While this question can't be answered, fishery-independent survey data for the deeper (> 30 m depth) segment of the uku stock should be included in future assessments to better understand trends in the relative abundance. Aspects of uncertainty pertaining to data were addressed in the response to TOR 1.The nineteen alternative model scenarios (i.e., sensitivity runs) in the assessment that addressed uncertainty in natural mortality, growth, maturity, recruitment variation, stock-recruitment steepness, recreational catch reconstruction, CPUE index combinations, and size frequency data likelihood weight typically only evaluated one source at a time and did not address multiple sources of uncertainty simultaneously. Thus, we're unable to address the multiplicative effects of multiple interacting sources of uncertainty on the assessment results. With that stated, the vast majority of runs (except for L_{Amax}), provided similar values for biological reference points and status determinations of not overfished and no overfishing status. The stability of these results may reflect a stock that is high on the stock-recruitment curve that is resilient to perturbations. Another explanation for the high stability may be the use of multiple fixed parameter values in the model (h, M, gear selectivity). For example, a series of sensitivity models demonstrated a considerable stability in model results across a range of scenarios. The fixed selectivities for some gears (determined outside of SS3 using LBSPR) may have contributed to this stability. The panel suggested the examination of temporal changes in catchability that the assessment demonstrated had little change on the model results (Fig. 7). There should be additional examination of catchability in future assessments to more accurate characterize how the fishing power of these gears/fleets has changed through time.

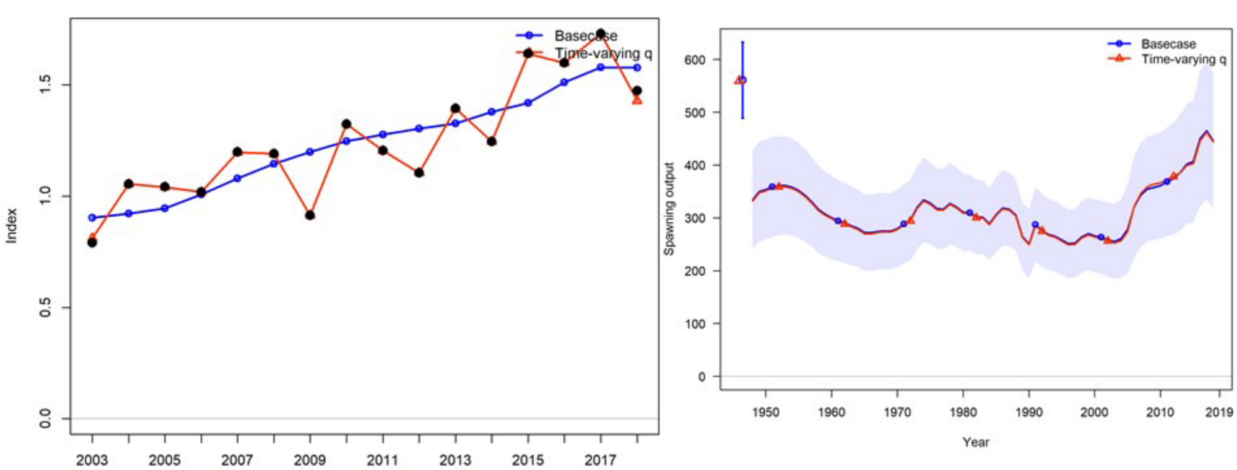

Figure 7. A comparison of fixed (base case) and time-varying catchability (via random walk) on deep-sea handline CPUE and SSB.

TOR 6 Assumptions: *Are model assumptions reasonably satisfied?*

Yes, the model assumptions were reasonably satisfied.

The model assumptions included fixed constant recruitment, steepness, CV in recruitment, natural mortality, catchability, and selectivity with the gears other than DSH estimated using the LBSPR method. There were assumptions inherent in the life history parameters used and input data which were well documented in the stock assessment report. Many of the assumptions were tested and reasonably satisfied in the sensitivity runs. To better organize the report assumptions, I recommend the addition of a table to include key biological and statistical assumptions to provide a single location for this information in the stock assessment report.

TOR 7 Scientific Validity: *Are the final results scientifically sound, including but not limited to estimated stock status in relation to the estimated overfishing and overfished status determination criteria (SDC)?*

Yes, the final results are scientifically sound, including but not limited to estimated stock status in relation to the estimated overfishing and overfished status determination criteria (SDC).

The base case assessment incorporates the best scientific information available and provides scientifically sound results that can be used to determine overfishing and overfished status for the main Hawaiian Islands uku stock. The stock assessment results find that the uku stock is not overfished, nor it is experiencing overfishing. These findings were robust under a variety of sensitivity runs that tested many aspects of the model inputs, uncertainties, and assumptions. The stability of these results suggests a high level of confidence in using these results for status determination. The reason that the results were so stable may derive from the consistency of patterns in the data used as indices, the relative high stock biomass on the stock-recruitment curve, and the use of fixed parameter values in the model. To facilitate the comparison of reference points generated by the base case and sensitivity runs, the final assessment report should contain a summary table.

TOR 8 Projection Methods: *Are the methods used to project future population state adequate, including the characterization of uncertainty, and appropriately applied for implementation of overfishing limits (OFL)?*

Yes, with caveats, the methods used to project future population state were adequate, properly included the characterization of uncertainty, and were appropriately applied for implementation of overfishing limits.

The age-structured model AGEPRO (Brodziak et al. 1998) was used for stock projections of uku for 2020-2026. The projections used base case model results from 2018 as initial conditions and 100 bootstrap replicates were calculated in SS to characterize population uncertainty. For each projection 1,000 simulations were run for each bootstrap replicate giving 100,000 simulated trajectories in each projection. The approach included three hypotheses about future recruitment levels: (1) similar to recent recruitment (i.e., 2000-2017; 20% of simulations), (2) similar to long term recruitment (i.e., 1949-2017; 20% of simulations), and (3) a Beverton-Holt recruitment curve with parameters from the SS3 base case model (60% of simulations). The rationale for using this particular mix of recruitment simulations was never clearly articulated and reflects the inherent uncertainty with predicting future events. The uncertainty contained in the recruitment aspect of the projects can not be understated especially given the potential influences of climate change on biological dynamics (which are not addressed in the models) so I recommend that the projections results only be presented to 2023, not 2026, in the final assessment. During the review, the panel requested an examination of the uncertainty in the projections related to recruitment and selectivity. From the results presented during the review, I recommend that the base case stock projections incorporate a composite selectivity for all gears weighted by catch contribution (DSH: 27%, ISH: 6%, Trol + Other: 12%, Recreational: 54%) with future catch determined from the average proportion of each fleets catch in 2016-2018, rather than a selectivity from only deep-sea handline. Given the uncertainty with predicting future recruitment and to acknowledge that recent recruitment has been historically high, I recommend that the assessment also include a worst case scenario that projects future scenarios using only values from the lowest 25th percentile of recruitment. While this is not suggested as an alternative for the base case, it provides managers with a possible worst case recruitment condition to visualize the effects on stock status.

TOR 9 Management: *Can the results be used to address management goals stated in the relevant FEP or other documents provided to the review panel? If any results of these models should not be applied for management purposes with or without minor short-term further analyses (in other words, if any responses to any parts of questions 1-8 are "no"), indicate:*

Which results should not be applied and describe why, and

Which alternative set of existing stock assessment results should be used to inform setting stock status and fishery catch limits instead and describe why.

Yes, the results can be used to address management goals stated in the relevant FEP.

The stock assessment included in the draft assessment report is the best scientific information available to make decisions relevant to the management of the main Hawaiian Islands uku stock and fishery. While none of my responses to TORs 1-8 were "no", there were some responses that were "yes with caveats" but the concerns raised from the caveats were sufficiently addressed during the review by sensitivity runs by the assessment team. These sensitivity runs demonstrated stable results with similar results and management recommendations to the base case. I suggest that many of the additional analyses requested should be incorporated into the final document (see TOR 10). In conclusion, the recommendations from the base case stock assessment are robust relative to uncertainties and assumptions from models and data, present reasonable decision points and parameterizations, and provide a feasible set of projections that generate management recommendations that can be used to address fishery management goals.

TOR 10 Recommendations: *As needed, suggest recommendations for future improvements and research priorities. Indicate whether each recommendation should be addressed in the short/immediate term (2 months), mid-term (3-5 years) and long-term (5-10 years). Also indicate whether each recommendation is high priority (likely most affecting results and/or interpretation), mid priority, or low priority.*

Short/immediate (2 months) term

All short/immediate term recommendations should be considered high priority and incorporated into the final version of the stock assessment.

To better represent the scope of the assessment, I suggest the alternative title, "Benchmark Stock Assessment of Uku (*Aprion virescens*) in the Main Hawaiian Islands, 2020".

To improve the reproducibility of the assessment results, add an Appendix to the assessment document or create a supplementary technical memo of all analysis code and data with guidance on how to reproduce results.

Figures to include in the final assessment:

(1) Time series of catch series by fleet/gear for the period 1948 to 2018
(2) Time series of fishing mortality time series by fleet/gear for 1948-2018
(3) Growth model that displays the samples from MHI and NWHI as separate colored dots
(4) Time series of standardized CPUE gear comparison (DSH, ISH, Troll) between the effort units of fishing day and fishing hours for 2003 to 2018
(5) Include the Mohn's rho value for biomass, recruitment and fishing mortality in retrospective figure
(6) Maps of diver survey locations for each survey year

Tables to include in the final assessment:

(1) List of biological and statistical assumptions for the assessment models which references appropriate sections of assessment report
(2) Summary table for management statistics for base-case and sensitivity runs that includes F_{2018}, F_{MSY}, $F_{2018/FMSY}$, SSB_{MSY}, SSB_{MSST}, SSB_{2018}, SSB_{2018}/SSB_{MSST}, and C_{MSY}. Use base-case and scenarios as rows in table. See list of sensitivity runs below.

Sensitivity runs to include:

(1) Use Lorenzen M in SS3 instead of a fixed natural mortality for all age classes
(2) Use a single DSH CPUE from 1948 to 2018 with same effort standardization (i.e. fishing days)
(3) Use extending ISH to 1992 to 2018 with same effort standardization (i.e. fishing days)
(4) Sensitivity run that combines all three of (1), (2), (3)
(5) Start model 1970 instead of 1948 to evaluate influence of additional catch data on assessment results
(6) Estimate SigmaR within the model rather than using a fixed value of sigma R
(7) Use iterative effective N (effN) estimation for data fitting

(8) Use a time-varying catchability for DSH from 2003 to 2018 to compare to results for a fixed q

Projections to include:

(1) Edit the base-case projection to use a composite selectivity for the different fleet/gear by their catches
(2) Include a worst-case projection with composite selectivity and 25th percentile lowest recruitment to provide an alternative scenario to consider under poor recruitment. This is not a replacement for the base-case.

Mid-term (3-5 years)

(1) Design and perform a simulation study to examine the potential sources of uncertainty in recreational fisheries data and identify an effective and optimal survey design. (HIGH)
(2) Continue with the direct involvement of MRIP survey statistician for stock assessments to assist with inclusion of highest quality recreational fishery data. (HIGH)
(3) Incorporate the fishery-independent survey results from the BFISH program as a relative index of abundance for uku. (HIGH)
(4) Communicate with fishermen active in the fishery to determine effective methods to improve the quality of non-commercial catch and effort information. (HIGH)
(5) Explore possible inclusion of the following elements for future base-case scenarios (HIGH)
    a. Using Lorenzen M
    b. Using a single DSH CPUE from 1948-2018 with same effort standardization
    c. Using extending ISH to 1992 to 2018
(6) Use the dealer reports to compare with the fisher reporting system records to improve for the quality of fishery-dependent catch data in future assessments (HIGH)
(7) Examine changing fleet dynamics with new technologies and gears to better estimate temporal variability of catchability, q. (MID)
(8) Further investment in life history with a focus on updating the reproductive and maturity studies that include an examination of sex ratio and fecundity-at-size. (MID)
(9) Continue collecting otolith samples for future growth studies. (LOW)

Long-term (5-10 years)

(1) Examine the dynamics and distribution of spawning aggregations in Hawaii. (HIGH)
(2) Better understand stock structure, population connectivity, and adult movement of fishes using genetic analysis and tagging experiments. (MID)
(3) Continue to improve data quality from fisheries-dependent and -independent sources. (HIGH)

TOR 11. Report: Draft a report (individual reports from each of the panel members and an additional Summary Report from Chair) addressing the above TOR questions

This document represents my individual reviewer report which addresses all the above TOR questions. Following the SOW, I also drafted a summary panel report with assistance from the panel members and presented those results to stakeholders on the last day of the review.

Synopsis on Recommendations and Consensus Summary

The stock assessment for the uku fishery in the main Hawaiian Islands concludes that the stock is not overfished, nor experiencing overfishing. Given the results presented in the assessment report and the additional information presented by the assessment team during the review, I recommend that the base case represents the best scientific information available and should be used for managing this fishery. The panel reached a similar consensus recommendation for the assessment. The multitude of sensitivity runs presented in the assessment and during the review showed that the results are robust to potential uncertainties in the data, models, and parameterizations. Interestingly, the results were extremely stable, which may be due to the relatively "healthy" status of the stock biomass and low fishing effort, although exploring other reasons that may explain this behavior should be pursued in future update assessments. As with many fisheries in the Hawaiian Islands, the need for higher quality information describing non-commercial catch and effort is vital, especially for this fishery which has a large relative catch and significant technological improvements in the fleet that focus on a summer spawning period on Penguin Bank. For the final version of the stock assessment report, several short/immediate term recommendations need to be incorporated that include figures and tables, sensitivity runs, and projections (see TOR 10 for details).

References

Craig MT et al. (2011) High genetic connectivity across the Indian and Pacific Oceans in the reef fish Myripristis berndti (Holocentridae). Marine Ecology Progress Series 334: 245-254.

Hordyk AR, Ono K, Prince JD, Walters CJ (2016) A simple length-structured model based on life history ratios and incorporating size-dependent selectivity: application to spawning potential ratios for data-poor stocks. CJFAS 73: 1787-1799.

Maunder MN, Sibert JR, Fonteneau A, Hampton J, Kleiber P, Harley SJ (2006) Interpreting catch per unit effort data to assess the status of individual stocks and communities. ICES Journal of Marine Science 63 (8): 1373–1385

Nadon MO (2017) Stock assessment of the coral reef fishes of Hawaii. U.S. Dep. Commerce, NOAA Tech. Memo., NOAA-TM-NMFS-PIFSC-60, 212 p.

Nadon MO, Sculley M, Carvalho F (2020) DRAFT Stock assessment of uku in Hawaii, 2020. NOAA Tech Memo., 105 p.

Sabater M, Kleiber P (2013) Improving specification of acceptable biological catches of data-poor reef fish stocks using a biomass-augmented catch-MSY approach. Western Pacific Regional Fishery Management Council, Honolulu, HI. 24 p.

Toonen RJ, et al. (2011) Defining boundaries for ecosystem-based management: a multispecies case study of marine connectivity across the Hawaiian Archiopelago. Journal of Marine Biology 2011: 1-13.

Appendix

Annex 2: Terms of Reference for the Peer Review

2020 Benchmark Stock Assessment for Main Hawaiian Islands Uku (*Aprion virescens*)

External Independent Peer Review under the Western Pacific Stock Assessment Review framework: 2020 Benchmark Stock Assessment for Main Hawaiian Islands Uku

For questions 1-8 and their subcomponents, reviewers shall provide a "yes" or "no" answer and will not provide an answer of "maybe". Only if necessary, caveats may be provided to these yes or no answers, but when provided they must be as specific as possible to provide direction and clarification to NMFS.

1. Of the data considered for inclusion in the assessment, were final decisions on inclusion/exclusion of particular data appropriate, justified, and well-documented?
2. Is the CPUE standardization properly applied and appropriate for this species, fishery, and available data?
3. Are the assessment models used reliable, properly applied, adequate, and appropriate for the species, fishery, and available data?
4. Are decision points and input parameters reasonably chosen?
5. Are primary sources of uncertainty documented and presented?
6. Are model assumptions reasonably satisfied?
7. Are the final results scientifically sound, including but not limited to estimated stock status in relation to the estimated overfishing and overfished status determination criteria (SDC)?
8. Are the methods used to project future population state adequate, including the characterization of uncertainty, and appropriately applied for implementation of overfishing limits (OFL)?
9. Can the results be used to address management goals stated in the relevant FEP or other documents provided to the review panel? If any results of these models should not be applied for management purposes with or without minor short-term further analyses (in other words, if any responses to any parts of questions 1-8 are "no"), indicate:

    Which results should not be applied and describe why, and

    Which alternative set of existing stock assessment results should be used to inform setting stock status and fishery catch limits instead and describe why.

10. As needed, suggest recommendations for future improvements and research priorities. Indicate whether each recommendation should be addressed in the short/immediate term (2 months), mid-term (3-5 years) and long-term (5-10 years). Also indicate whether each recommendation is high priority (likely most affecting results and/or interpretation), mid priority, or low priority.
11. Draft a report (individual reports from each of the panel members and an additional Summary Report from Chair) addressing the above TOR questions.

Panel Meeting Agenda and Participants

WPSAR panel: Chair Erik Franklin (WPRFMC SSC and University of Hawaii), Yong Chen (University of Maine), Yan Jiao (Virginia Tech)

WPSAR Coordinating Committee: Marlowe Sabater (WPRFMC), Brett Schumacher (NOAA PIRO), John Syslo (NOAA PIFSC)

Stock Assessment Team: Marc Nadon (NOAA PIFSC), Michelle Sculley (NOAA PIFSC), Felipe Carvalho (NOAA PIFSC)

Attendees: Mike Seki (NOAA PIFSC), Joe O'Malley (NOAA PIFSC), Hongguang Ma (NOAA PIFSC), Roy Morioka (public, fisher), Bryan Ishida (Hawaii DAR), Beth Lumsden (NOAA PIFSC), T. Todd Jones (NOAA PIFSC)

<u>Day 1, Monday February 24</u>

1. Welcome and Introductions (Syslo)
2. Background information – Objectives and Terms of Reference (Syslo)
    a. Fishery Operation (Morioka)
    b. Fishery Management (Schumacher)
3. History of stock assessments and reviews (Carvalho)
4. Data (Nadon)
    a. Hawaii Division of Aquatic Resources Fishing Report System (FRS) and Hawaii Marine Recreational Fishery Survey (HMRFS)
    b. Life history information
    c. Other
5. Presentation and review of stock assessment (Nadon)

<u>Day 2, Tuesday February 25</u>

6. Continue presentation and review of stock assessment (Nadon, Sculley, Carvalho)

<u>Day 3, Wednesday February 26</u>

7. Continue review of stock assessment (Nadon, Sculley, Carvalho)
8. Panel discussions (closed)

Day 4, Thursday February 27

9. Continue review of stock assessment (Nadon, Sculley, Carvalho)
10. Public comment period
11. Panel discussions (closed)

Day 5, Friday February 28

12. Present panel summary results (morning)
13. Adjourn