An Independent Peer Review of the 2023 Benchmark Stock Assessment for the American Samoa Bottomfish February 17-23, 2023 Pago Pago, American Samoa

Submitted to the
Center for Independent Experts

by
Joseph E Powers

Executive Summary

A 2023 benchmark assessment was conducted implementing the first integrated single species stock assessments for the American Samoa bottomfish. These integrated assessments used the Stock Synthesis framework to incorporate CPUE indices, size frequency, and catch data into a single age-structured model for nine of the 11 American Samoa Bottomfish Management Unit Species. These integrated models were used to estimate biomass and stock status through time, and stock status was evaluated against MSY-based reference points described in the Fishery Ecosystem Plan for the American Samoa Archipelago. Projections were provided to inform management and the setting of annual catch limits. The results of the nine analyses indicated that all current spawning stock biomasses are greater than SSBMSY and that all current fishing mortality rates are less than FMSY. A large number of sensitivities were explored prior to and at the review meeting. The results of which supported the assessment conclusions. Recommendations for future assessment work include: 1) re-examining the inclusion of 1986-2015 CPUEs in the assessments; 2) consider recruitment deviations; 3) continued examination of life history parameters; and 4) consider selectivity changes over time. Additionally, two species were not assessed with these models due to inadequacies of data. Based on FEP criterion, those two species were each assigned an “indicator species” for further management action. Those assignments were based on taxonomy, life history and habitat.

Background

A stock assessment benchmark of the American Samoa bottomfish was conducted through 2021 by PIFSC scientists. The deep-slope fishes of American Samoa support a small yet valuable boat-based fishery in depths ranging around 100 m to 400 m. The Western Pacific Regional Fishery Management Council's fishery ecosystem plan for American Samoa includes 11 bottomfish management unit species (BMUS) that have traditionally been assessed and managed as a species complex. The previous assessment (Langseth et al. 2019) evaluated the BMUS as an aggregate using a Bayesian stock production model and provided management advice for that aggregate rather than for the individual species of the BMUS. At the time it was noted that the aggregate results would not allow detailed species recommendations. Thus, subsequently a concerted effort was made to disaggregate the data (catches, CPUEs and size sampling) by species and move toward species-specific assessments, To that end the current American Samoa benchmark assessment split the complex into its component BMUS (i.e. single-species assessments). This allowed the implementation of age-structured models and the incorporation of size and life history information. Moving assessment models from a complex-level to a finer taxonomic resolution meant that special consideration needed to be given to species identification in the various data sets. These considerations followed recommendations from a series of community workshops that involved fishers, managers, and scientists on best practices for analyzing bottomfish catch and effort, and size data for use in stock assessments. The 2023 benchmark assessment diverges significantly from this previous work as it implements the first integrated single species stock assessments for the American Samoa bottomfish. These integrated assessments use the Stock Synthesis framework to incorporate CPUE indices, size frequency, and catch data into a single age-structured model for each one of nine of the 11 species. These integrated models were used to estimate biomass and stock status through time, and stock status was evaluated against MSY-based reference points described in the Fishery Ecosystem Plan for the American Samoa Archipelago. Projections were provided to inform management and the setting of annual catch limits.

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 an independent peer review in accordance with the requirements specified in this Performance Work Statement (Appendix 2) and TORs, to assist the Peer Review Panel (co)Chairs 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 TORs and my overall conclusions follow.

Summary of Findings for Each Term of Reference

Overall Discussion of the Assessments

The current (2023) stock assessment of American Samoa (AM) bottomfish established markedly different assessment methods from that utilized in the previous assessment in 2019. The 2019 assessment addressed 11 Bottomfish Management Unit Species (BMUSs) with an aggregate biomass model (Pella-Tomlinson production model) fit with Bayesian methods to an aggregate CPUE. The catch history estimated in 2019 was the aggregate of the 11 BMUSs (Table 1). Conversely, the 2023 assessment moved to single species models for nine of the 11 species implemented in Stock Synthesis 3 (SS3). This action was in response to recommendations in 2019 to move toward single-species assessments. Additionally, the 2019 assessment results found that the aggregate biomass was below the Minimum Stock Size Threshold (MSST: established in the FEP as 70% of BMSY). This added impetus to reevaluation.

Table 1: 11 Bottomfish Management Unit Species

BMUSs ECTA and PRFI were not included in 2023 single species assessments		
Aphareus rutilans	Rusty jobfish	APRU
Aprion virescens	Green jobfish	APVI
Caranx lugubris	Black jack	CALU
Etelis carbunculus	Deep-water red snapper	ETCA
Etelis curuscans	Deep-water longtail red snapper	ETCO
Lethrinus rubrioperculatus	Spotcheek emperor	LERU
Lutjanus kasmira	Bluestripe snapper	LUKA
Pristipomoides filamentosus	crimson jobfish	PRFI
Pristipomoides flavipinnis	Golden eye jobfish	PRFL
Pristipomoides zonatus	Oblique-banded snapper	PRZO
Variola louti	yellow-edge lyretail grouper	VALO

The PIFSC analysts made this transition with a lengthy and thorough review process to recover data streams; review existing data and parse these aggregate data into species; and implement recent data obtained since the 2019 assessment. Specifically, this included dividing the 1986-2015 catch data (used as an aggregate in the 2019 assessment) into catches by species, recover from previously unknown records the 1967-86 catches and divide them into species, divide 1986-2015 CPUE data into species and standardize each, implement newly acquired species catches and standardized CPUEs (2016-2021) and utilize size frequency samples by species collected sporadically since approximately 2000. Additionally, the review noted a few small errors in the catch series previously used and corrected them.

Note that all eleven of these species are considered *data-poor*. What this means in the American Samoa context is that the catch and effort series by species are mostly reconstructed. Life history parameters (growth, natural mortality, etc) are not typically available for the eleven AM species and parameters are drawn from like-species in AM and from the same species (or like species) in other regions. The logical conclusion to this is that data-poor assessments in general are expected to be very uncertain and that future data-collection in AM will show the 2019 and 2023 assessments to be imprecise and probably

inaccurate. Nevertheless, the results provide an adequate basis for assessment and management going forward.

The general strategy for the nine 2023 single species assessments was to include species-specific catches 1967-2021, species-specific standardized Late CPUEs 2016-2021 and then estimate R0 (equilibrium recruitment with no fishing), catchability coefficient and selectivity parameters within SS3. Steepness was fixed, as were M and growth parameters. The growth parameters were determined from stepwise methods external to SS3. The motels were fitted to the late CPUE data and size frequencies for a single "fishery". The Beverton-Holt stock-recruitment model was used without annual deviations. Multiple sensitivities were explored in the assessment document by perturbing the above parameters and structural assumptions.

Much of the Feb 17-23 WPSAR review meeting was utilized in expanding the sensitivity analyses to explore uncertainties in parameter specifications. Additionally, analyses at the meeting attempted to provide a "bridging" between 2019 and 2023 assessments. i.e. what were the decision processes for data selection and what were the consequences. This was lacking in the original 2023 assessment document. To that end I cumulated the aggregate of the nine single-species results from the 2023 assessment and compared those results with the 2019 assessment which was an aggregate of 11 species, Additionally, the species catch and composition are given (following Figures 1 - 3):

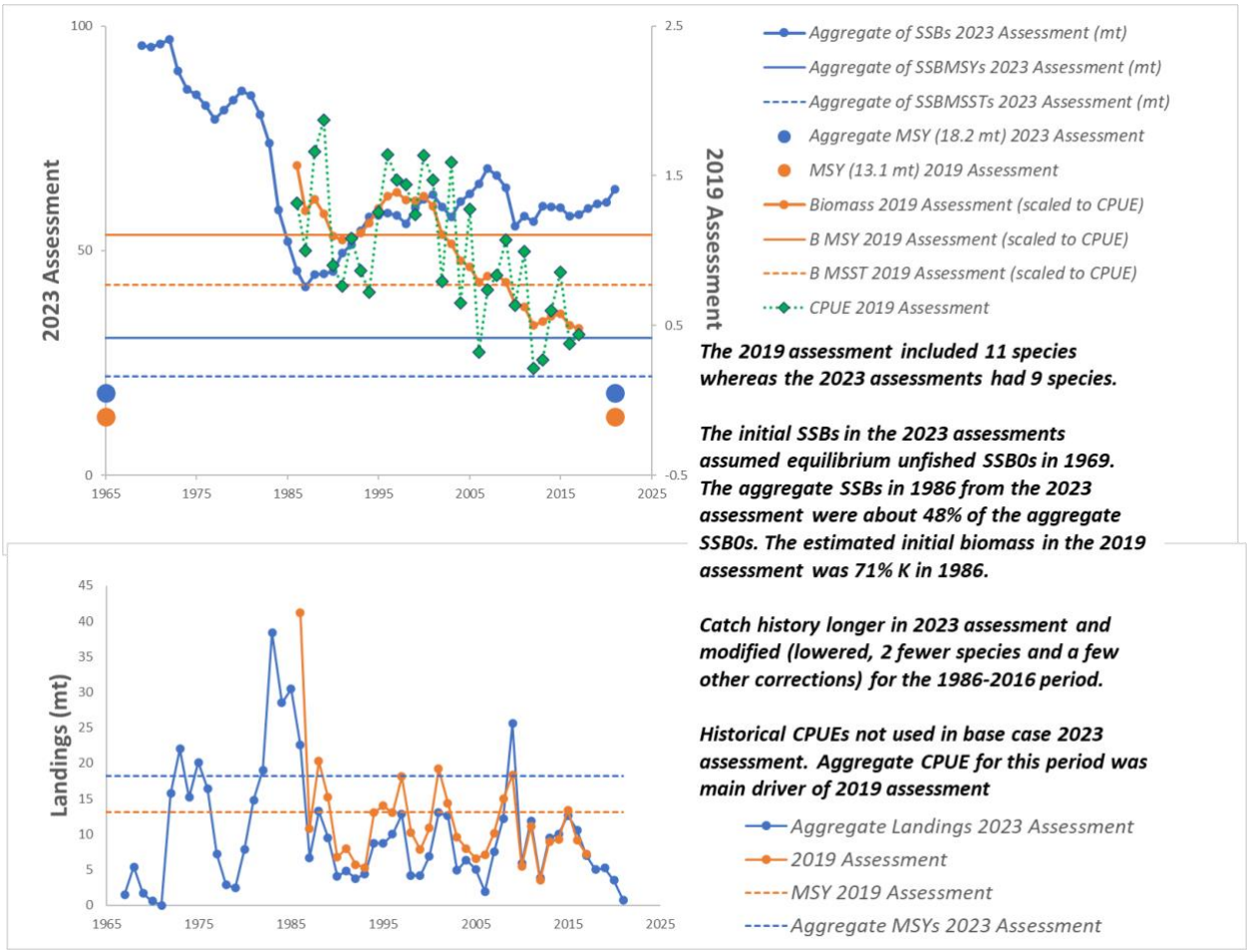

Figure 1 : Species assessments and aggregate landings

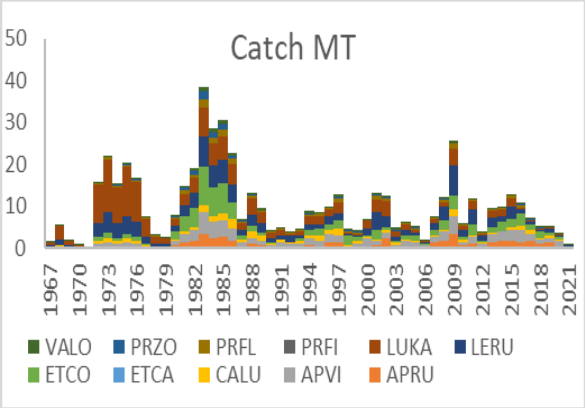

Figure 2: Species-based landings

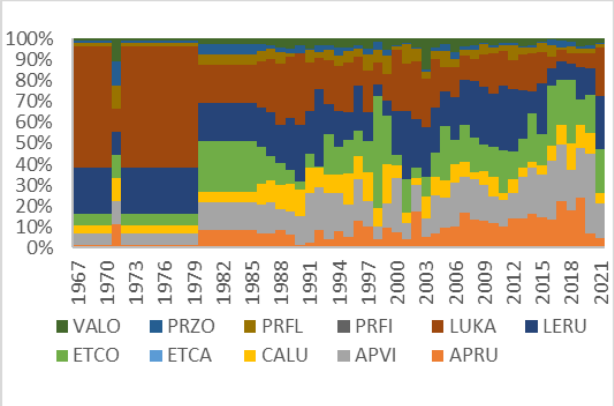

Figure 3: Species landings by proportion of total catch

There are a number of issues in comparing the 2019 and 2023 results even in the aggregate.

1) the aggregate catches 1986-2015 are not significantly different between the two assessments, given one has nine species and the other has 11. However, the species composition shifts quite a bit over time. Thus, I suspect the species shifts were not well handled in the 2019 CPUE standardization. For example LUKA: its catches went down during this period, yet it is one of the higher biomass species in the 2023 assessments. The early (1986-2015) species-specific CPUEs (standardized) data points were compared with the aggregate CPUE data points of the 2019 assessment in terms of linear trend. This is a convenient simple demonstration of the differing trends in the CPUE data and how the implicit weighting in the 2019 aggregate favored the declining trend of LUKA, especially.

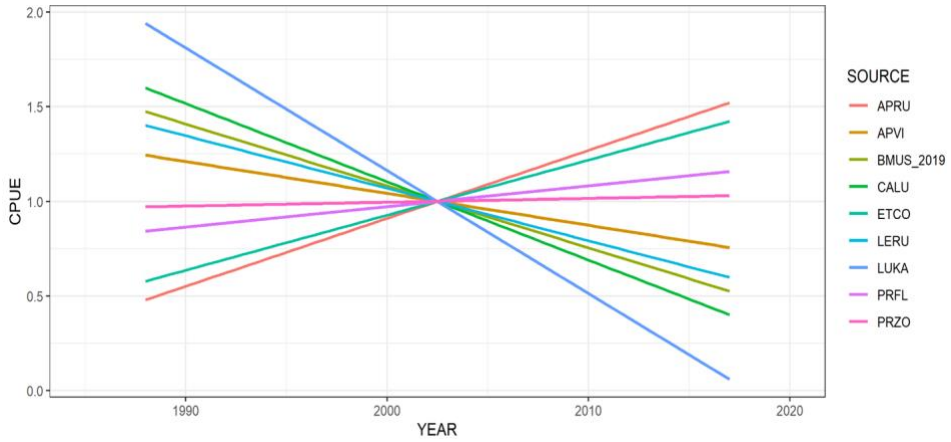

Figure 4: Trends in CPUE data

2) Catch estimates for 1967-1985 were not available for the 2019 assessment. The 2018 analysts attempted to compensate for this by allowing for the 1986 biomass to be less than K (carrying capacity). Their estimate was B1986/K~71%. Conversely, the 2023 assessment assumed that SSB0 (equivalent to carrying capacity in SSB units) existed in 1966 and then each biomass was affected by that species' subsequent catch. The aggregate of the SSBs in 1986 from the 2023 assessment was about 48% of the aggregate SSB0s. I believe the prior distribution for B86/K used in 2019 did not include 48%. Thus, the 2023 assessment included early catches that exceeded those expected by the 2019 model.

3) The definition of MSST for single species management is SSBmsy*(1-M) or SSB*0.5 whichever is greater. For this FEP, the generic MSST for aggregate species management is Bmsy*(1-0.3). In the 2019 assessment, the Pella Tomlinson results yielded an exponent of 2.66 in the model, thus the results were Bmsy/K=0.55
and MSST/K=0.55*0.7=38.5%. Conversely, the 2023 results for most species were SSBmsy/SSB0(equiv to carrying capacity) of about 25-35%. Thus, with (1-M)'s in the range of 0.9 to 0.5, the MSST/Ks are a lower criteria in the single species framework.

4) The premise of most aggregate assessments and aggregate management units is that the life history of the constituent species is "similar" (Zhou et al. 2019. Rev Fish Biol Fisheries 29:711-733). The Ms used in the 2023 assessment suggest two or perhaps three very different categorizations of productivity.

Species	ETCO	APVI	APRU	PRZO	PRFL	LERU	VALO	CALU	LUKA
M	0.1	0.17	0.18	0.18	0.19	0.36	0.36	0.45	0.68

If in fact the M's vary this much among the nine species, it is not surprising that the aggregate dynamics are being masked. I cannot determine if this diversity in productivity indicators (Ms) was noted in the 2019 assessment, nor whether the high M's used in 2023 for LUKA especially was a change in scientific opinion from 2019.

5) Aggregate catches in the early 1980s were in excess of 30 mt and recent catches have been small especially in the COVID years. It is interesting to note that the aggregate of the nine MSYs from the 2023 assessment is 18.2 mt whereas the MSY for the 11 species aggregate in 2019 was 13.1 mt. This outcome lines up with the discussions above. The expectation from recent low catches is that each stock is currently unlikely to be heavily impacted. The assessments are consistent with that.

All of this is more of a comment on aggregate assessments and management than the current assessment methods and results. Aggregate data can mask trends in individual species and that appears to have happened in this case. Additional data streams (especially early catches) added to the understanding of the species dynamics and but the magnitudes were not anticipated to that degree in the 2019 assessment. The aggregate status criterion was more restrictive than the individual species SDCs, as well. The experience denoted is meant to provide guidance for aggregate management and assessment in data-poor situations.

Ideally, the analysts should have provided an update of the 2019 aggregate assessment model using the 2013 aggregated data to assist with the "bridging" process of understanding what changes are model driven versus data driven. In this circumstance, the choice of how one aggregates current data and implements current CPUEs to apply to the old model is not trivial. But, as a general recommendation, assessments should include this phase

The current 2023 assessments implemented standardized CPUE series for each species for the 2016-2021 period. The analysts felt that the data collection scheme was sufficiently different such that combining with the earlier data was not justified. Thus, they were left with two CPUE series for each species early (~1987-2015) and late (2015-2021). However, the late series for each species had few data points (years), with a relatively high std dev and typically without much detectable trend. Therefore, in a practical sense, these data did not have much leverage in the model fitting process. Including or not including that series typically did not change the outcomes much. The analysts chose not to include the early standardized CPUEs in the base cases. It is unclear the criteria for not including them; perhaps, lack of confidence in the species separation methods of the CPUE data. However, the final effect was that fits were largely driven by the size frequency data. Again, including vs not including both early and late CPUE data were explored through sensitivity analyses with key species and the net effect was marginal, especially in terms of status determination criteria.

Generally, my "priors" for an assessment are 1) to include data series (such as CPUEs) that were used in the previous assessment unless there are compelling reasons not to; 2) to include recruitment deviations in the years where they can be estimated. My reasons are that dynamics are most often driven by year-class strength. Also, much of our perception of the probabilities of achieving status criteria are related to "process error", i.e., recruitment deviations (sigmaR), as well as the precision estimates in the assessment itself; and 3) to be somewhat skeptical of flat-topped selectivities, especially where the perceived selectivity is driven by spatial expansion or contraction of the fishery.

Having said this, I can be convinced to modify my priors based the existing situation. In this case the 1986-2015 data sets used in 2023 were not exactly what was used in 2019. Additional, methods (and assumptions) were implemented to obtain species-specific CPUEs. Thus, the "data" are not exactly the same. Unless an argument is presented to suggest otherwise, I still recommend incorporating early CPUE data into the single-species assessments in the final assessment report as sensitivities or as base cases (if warranted). In either case, I don't expect the SDC decision outcomes to change.

I recognize that in these data-poor situations there is little expectation of estimating very precisely annual recruitment deviations. Perhaps, the best that can be done with the data at hand is to estimate recent trends and their scale such that relevant catch quantities (MSY, ABC, ACT, etc) can be estimated. Thus, in this case I can accept not including recruitment deviations. However, the expected underestimation of variance in projections and SDC probabilities should be noted.

Similarly, under these data-limitations I can accept the flat-topped model structure at this time. A sensitivity run using dome-shaped selectivity throughout the time series showed that SDC decision outcomes were unchanged, median trends were very similar, but the scale of the biomass increased. This was expected. Note that the SDC's are usually based on ratios (SSB/SSBmsy or SSB/SSB0). Therefore, the change in scale induced by dome selectivities changes both the numerator and denominator of the SDC ratios, making them less sensitive to changes in selectivity. Nevertheless, my feelings are that it is more likely that there has been a shift in selectivity from the early years to now, but this is just conjecture. If so, the trends might change but recent trends and scale will be less affected. Thus, with the current data and the focus being on the recent status, the current selectivity pattern is accepted.

I, like others, at the meeting focused on LUKA. There are several issues: 1) M=0.68 (Amax=8) and h=0.69 seems incompatible; 2) the resulting R/SSB from the assessment has implications about the biology of LUKA compared to other species; 3) model estimates of F in several early years appeared to not achieve the observed catch; and 4) the declining trend in early CPUE is not seen in the assessment to the same degree. Despite these issues it is very unlikely that LUKA SDC decisions will change from what is presently in the assessment.

With LUKA M=0.68 (derived from an Amax of 8), I would expect a more herring-like life history strategy with h of around 0.9 or more and sigmaR's of 0.5 or more and an SSBMSY/SSB0 about 20% or less. The choice of h=0.69 appears to be based on a general Lutjanid life history (Thorson 2019), but with M=0.68, is LUKA like a general Lutjanid? Choosing h=0.69 puts SSBMSY/SSB0 at about 40%. In this case, the h-choice doesn't have a large impact on the assessment fits because the observed biomass appears to limit the recruits to the flat-topped part of the S-R curve, regardless of whether h is 0.9 or 0.69. But it does have an impact on SSBMSY/SSB0. Fortunately, if SSB status criteria are met with h=0.69, they will be met with h=0.9.

With LUKA, the implications of M=0.68 results in model estimates of R/SSB being very different from the R/SSB of the other species. The biological implication is that survival in the first year (recruitment) is very different compared to the other species or that the number of eggs produced by a kg of SSB is very different or both. But is that expected from this Lutjanid? According to the Thorson (2019) meta-analysis approach, the answer is no. But I leave it to species experts.

Examination of the detail of LUKA model runs at the meeting showed that in several early year catches, the estimated F's appeared to hit the estimation boundary in SS3 and thus the predicted catches at that F were less than the observed. This, needs to be examined more thoroughly. One alternative going into the future is to acknowledge that catches are estimated with error, especially in the early part of the time series. Then F's are estimated based on observed catches with an appropriate error structure and appropriately weighted in the MLE process. This can be implemented within the current SS3 structure.

Given the impact LUKA had on the early aggregated CPUE, I'd like to see the implementation of the early LUKA CPUE into the assessment explored more thoroughly.

In this data-poor situation an overarching objective is to provide an assessment-data template going forward as new data are collected. These assessments have done that. Future assessments will have continuity with previous assessments with known changes in data and model structure.

Given this discussion, my short specific responses to the specific TORs are as follows:

Response to each Term of Reference

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

Yes.

Note that the 1987-2015 CPUE time series was excluded without sufficient justification for the base models in the assessment report. This was examined at the meeting and results suggested that it did not impact the status determinations decisions. Nevertheless, the justification process should be elaborated in the final assessment report.

Is the CPUE standardization properly applied and appropriate for this species, fishery, and available data?

Yes.

CPUE standardization methodology was properly applied and appropriate for all assessments based on available data.

Are the assessment models used reliable, properly applied, adequate, and appropriate for the species, fishery, and available data?

Yes.

The assessment models using SS3 were reliable, properly applied, adequate, and appropriate for these species based on available data.

Are decision points and input parameters reasonably chosen?

Yes.

The decision points and input parameters were reasonably chosen.

Bridging analysis is needed to justify shift from a species complex model in 2019 to single species models in 2023 (completed during review).

Most life history parameters were not from local studies.

Choice to fix many of the input parameters led to an underestimate of the uncertainty in assessment results but these concerns were adequately explored in sensitivity analyses.

Are primary sources of uncertainty documented and presented?

Yes.

The primary sources of uncertainty were documented and implemented to the extent feasible with the data at hand. Additionally, uncertainty in life-history parameters were evaluated extensively including at the review meeting. Nevertheless, future assessments should examine impacts of recruitment variability and time-variation in selectivity on overall uncertainty.

Are model assumptions reasonably satisfied?

Yes.

Per the discussion above, the key model assumptions for these assessments were that life history parameters known, that selectivity was flat-topped, that length frequencies were representative of the catch and that the observed catch known with limited error. These assumptions were evaluated in the assessment report and at the meeting. Thus, at the current stage of model/data these assumptions are reasonably satisfied.

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.

As noted in the discussion above, many additional sensitivities were explored at the meeting. Those, additional sensitivity runs performed during the review should be added to the assessment to better inform managers of the uncertainty in the results

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.

Per the discussion above the assessments are necessarily based on fixed parameters, uncertain catch histories and no process error required of data-poor stock assessments. The obvious result is that true uncertainty is underestimated.

This probably doesn't affect the OFL, as that is the median of the uncertainty distribution. However, it does affect the overall distribution which has implications for ABC and ACT as determined by the SSC and the Council. The final assessment report should discuss this.

If applied, is the choice of indicator species to evaluate more poorly known species that are in a stock complex appropriate?

Yes.

Two species were not assessed with the models (ETCA and PRFI) and per the FEP require indicator species to be defined for each. Analyses presented at the meeting suggested that ETCO be used as an indicator species for ETCA and PRFL be used as an indicator species for PRFI. These were based on the indicator being of the same genus, similar maximum age, similar habitats (depth range) and similar length. These are appropriate.

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 1-9 are "no"), indicate which alternative option should be used to inform setting stock status and fishery catch limits between 1) using the previous assessment, 2) using an indicator species, or 3) designing the stock status as "unknown".

Yes.

Note that: 1) the current fishing effort is relatively low so overfishing is not a concern; 2) life history parameters and other issues of Lutjanus kasmira were noted above and should be checked for the final report; 3) stock status and fishery catch limits of indicator species should be used for the two species ETCA and PRFI as noted in the TOR above.

As needed, suggest recommendations for future improvements and research priorities. Indicate whether each recommendation should be addressed in the short/immediate term (for this assessment), mid-term (next assessment) 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.

<u>Short/immediate term (this assessment)</u>

- Incorporate a text section, model runs, and analyses to bridge between prior and current assessment data, model, and results per the guidance provided in the discussion above.
- For LUKA, have a stock assessment run with the historical catch series assumed to be measured with error (with appropriate CV and error structure).
- Incorporate life history sensitivity runs and analysis.
- Add section on indicator species. Present standardized criteria and justification for indicator species chosen.
- Add to report a summary table of biological reference points for each species

<u>Mid-term term (next assessment)</u>

- Perform local life history studies for BMUS species and utilize those parameters for base case assessments, with LUKA as a priority.

- Ensure standardized methods of creel survey provide reliable and representative catch, CPUE, and length composition data.
- Revisit incorporation of longer CPUE (1988-2015) time series in base models.
- Incorporate the additional sensitivity runs (Linf, M, selectivity) prepared for this assessment as a standard part of model evaluation.

Long term (5-10 years)

- Continue to perform local life history studies for BMUS species and utilize those parameters for base case assessments.
- Perform a fishery-independent survey to estimate BMUS density, abundance, biomass, and length composition.

Conclusions and Recommendations

The WPSAR Review meeting provided an excellent forum for interaction with PIFSC analysts and others at the meeting. This was an improvement from ZOOM meetings required during the COVID era. The cooperation and support of the PIFSC and PIRO analysts and staff, DMWR and the Chair were exemplary.

In this data-poor situation an overarching objective is to provide an assessment-data template going forward as new data are collected. These assessments have done that. Future assessments will have continuity with previous assessments with known changes in data and model structure. The results of the nine analyses indicated that all current spawning stock biomasses are greater than SSBMSY and that all current fishing mortality rates are less than FMSY. A large number of sensitivities were explored prior to and at the review meeting. The results of which supported the assessment conclusions.

A number of recommendations are suggested: 1) re-examining the inclusion of 1986-2015 CPUEs in the assessments; 2) consider recruitment deviations; 3) continued examination of life history parameters; and 4) consider selectivity changes over time; 5) building a better “bridge” between current and previous assessments; and 6) reexamine LUKA inputs and model structure.

Appendix 1: Bibliography of materials provided for review

Carvalho, F., Winker, H., Courtney, D., Kapur, M., Kell, L., Cardinale, M., Schirripa, M., Kitakado, T., Yemane, D., Piner, K.R. and Maunder, M.N., 2021. A cookbook for using model diagnostics in integrated stock assessments. Fisheries Research, 240, p.105959.

Erickson, K.A. and Nadon, M.O., 2021. An extension of the stepwise stochastic simulation approach for estimating distributions of missing life history parameter values for sharks, groupers, and other taxa. Fishery Bulletin, 119(1), pp.77-114.

Langseth B, Syslo J, Yau A, Carvalho F. 2019. Stock assessments of the bottomfish management unit species of Guam, the Commonwealth of the Northern Mariana Islands, and American Samoa, 2019. NOAA Tech Memo. NMFS-PIFSC-86, 177 p. (+ supplement, 165 p.). doi:10.25923/bz8b-ng72.

Ma, H., Matthews, T., Nadon, M., Carvalho, F. 2022.Shore-based and boat-based fishing surveys in Guam, the CNMI, and American Samoa: survey design, expansion algorithm, and a case study. U.S. Dept. of Commerce, NOAA Technical Memorandum NOAA-TM-NMFS-PIFSC-126, 115 p. doi: 10.25923/c9hn-5m88

Methot Jr, R.D. and Wetzel, C.R., 2013. Stock synthesis: a biological and statistical framework for fish stock assessment and fishery management. Fisheries Research, 142, pp.86-99.

Nadon, M.O. and Ault, J.S., 2016. A stepwise stochastic simulation approach to estimate life history parameters for data-poor fisheries. Canadian Journal of Fisheries and Aquatic Sciences, 73(12), pp.1874-1884.

Nadon MO, Bohaboy EC. 2022. Evaluation of the data available for bottomfish stock assessments in American Samoa. U.S. Dept. of Commerce, NOAA Technical Memorandum NMFS-PIFSC-123, 85 p. doi:10.25923/m74z-7d07.

Rudd, M.B., Cope, J.M., Wetzel, C.R. and Hastie, J., 2021. Catch and length models in the stock synthesis framework: expanded application to data-moderate stocks. Frontiers in Marine Science, 8, p.663554.

Thorson, J.T., 2020. Predicting recruitment density dependence and intrinsic growth rate for all fishes worldwide using a data-integrated life-history model. Fish and Fisheries, 21(2), pp.237-251.

Thorson, J.T., Johnson, K.F., Methot, R.D. and Taylor, I.G., 2017. Model-based estimates of effective sample size in stock assessment models using the Dirichlet-multinomial distribution. Fisheries Research, 192, pp.84-93.

Thorson, J.T., Bryan, M.D., Hulson, P.J.F., Xu, H. and Punt, A.E., 2020. Simulation testing a new multi-stage process to measure the effect of increased sampling effort on effective sample size for age and length data. ICES Journal of Marine Science, 77(5), pp.1728-1737.

Walter, J. and Winker, H., 2020. Projections to create Kobe 2 strategy matrix using the multivariate log-normal approximation for Atlantic yellowfin tuna. Collect. Vol. Sci. Pap. ICCAT, 76(6), pp.725-739.

Winker, H., Kerwath, S.E. and Attwood, C.G., 2013. Comparison of two approaches to standardize catch-per-unit-effort for targeting behaviour in a multispecies hand-line fishery. Fisheries Research, 139, pp.118-131.

Winker, H., Walter, J., Cardinale, M. and Fu, D., 2019. A multivariate lognormal Monte-Carlo approach for estimating structural uncertainty about the stock status and future projections for Indian Ocean Yellowfin tuna. 22nd Session of the IOTC Working Party on Temperate Tunas. IOTC-2019-WPTT21-51. Mahe, Seychelles: IOTC.

Yau, A., M. O. Nadon, B. L. Richards, J. Brodziak, and E. Fletcher. 2016. R Stock Assessment Updates of the Bottomfish Management Unit species of American Samoa, the Commonwealth of the Northern Mariana Islands, and Guam in 2015 Using Data through 2013. U.S. Dep. Commer. NOAA Tech. Memo., NOAA-TM-NMFS-PIFSC-51, 54p.

Appendix 2

Performance Work Statement for

Center for Independent Experts' Contribution of Reviewers to the Western Pacific Stock Assessment Review of the 2023 Benchmark Stock Assessment for the American Samoa Bottomfish

February 17-23, 2023

Background

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

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

Scope

A stock assessment benchmark of the American Samoa bottomfish was conducted through 2021 by PIFSC scientists. The deep-slope fishes of American Samoa support a small yet valuable boat-based fishery in depths ranging around 100 m to 400 m. The Western Pacific Regional Fishery Management Council's fishery ecosystem plan for American Samoa includes 11 bottomfish management unit species (BMUS) that have traditionally been assessed and managed as a species complex. The current benchmark assessment split the complex into its component BMUS (i.e. single-species assessments). This allowed the implementation of age-structured models and the incorporation of size and life history information. Moving

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

assessment models from a complex-level to a finer taxonomic resolution meant that special consideration needed to be given to species identification in the various data sets. These considerations followed recommendations from a series of community workshops that involved fishers, managers, and scientists on best practices for analyzing bottomfish catch and effort, and size data for use in stock assessments. The 2023 benchmark assessment diverges significantly from this previous work as it implements the first integrated single species stock assessments for the American Samoa bottomfish. These integrated assessments use the Stock Synthesis framework to incorporate CPUE indices, size frequency, and catch data into a single age-structured model for each one of the 11 species. These integrated models were used to estimate biomass and stock status through time, and stock status was evaluated against MSY-based reference points described in the Fishery Ecosystem Plan for the American Samoa Archipelago. Projections were provided to inform management and the setting of annual catch limits. The specified format and contents of the individual peer review reports are found in **Annex 1**. The Terms of Reference (ToRs) of the peer review are listed in **Annex 2**. Lastly, the tentative agenda of the panel review meeting is attached in **Annex 3**.

Requirements for Center of Independent Experts (CIE) Reviewers
NMFS requires two reviewers who are external to PIFSC, Pacific Islands Regional Office (PIRO), and the Western Pacific Regional Fishery Management Council and its affiliated bodies to conduct an impartial and independent peer review in accordance with this Performance Work Statement (PWS), OMB Guidelines, and the ToRs in Annex 2.

CIE reviewers shall have:

- Working knowledge and recent experience in the application of stock assessment models for data-limited and moderate fisheries, sufficient to complete a thorough review.
- Knowledge of integrated assessment models, more specifically Stock Synthesis;
- Expertise with measures of model diagnostics, uncertainty, forecasting, and biological reference points;
- Familiarity with federal fisheries science requirements under the Magnuson-Stevens Fishery Conservation and Management Act;
- Familiarity with local Pacific Islands fisheries as well as artisanal fisheries and fishing practices;
- Excellent oral and written communication skills to facilitate the discussion and communication of results.

The chair, who is in addition to the two reviewers, will be not be provided by the CIE. Although the chair will be participating in this review, the chair's participation (i.e. labor and travel) is not covered by this contract.

Tasks for Reviewers

Pre-review Background Documents

Each of the CIE reviewers shall complete the following tasks in accordance with the PWS and Schedule of Milestones and Deliverables.

Pre-review Background Documents: No later than two weeks before the peer review, the NMFS Project Contact will provide reviewers the necessary background information and reports for the peer review. The reviewers shall read all documents prior to the peer review in accordance with the PWS scheduled deadlines.

Required pre-review documents:

- DRAFT 2023 American Samoa Bottomfish Stock Assessment Report
- Previous 2019 bottomfish stock assessment: Langseth B, Syslo J, Yau A, Carvalho F. 2019. Stock Assessments of the Bottomfish Management Unit Species of Guam, the Commonwealth of the Northern Mariana Islands, and American Samoa (focus on sections pertaining to American Samoa)
- American Samoa Archipelago Ecosystem Plan: Western Pacific Regional Fishery Management Council. 2009.
- 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.
- Rudd, MB, Cope, JM, Wetzel, CH, and Hastie, J. 2021. Catch and length models in the stock synthesis framework: expanded application to data-moderate stocks. Frontiers in Marine Science.
- Nadon, MO and Ault, J. 2016. A stepwise stochastic simulation approach to estimate life history parameters for data-poor fisheries. CJFAS 73:1874-1884.

Panel Review Meeting

Each CIE reviewer shall conduct the independent peer review in accordance with the PWS and ToRs, and shall not serve in any other role unless specified herein. Each CIE reviewer shall actively participate in a professional and respectful manner as a member of the meeting review panel, and their peer review tasks shall be focused on the ToRs as specified herein. The meeting will consist of presentations by NOAA and other scientists to facilitate the review, to provide any additional information required by the reviewers, and to answer any questions from reviewers.

Contract Deliverables - Independent CIE Peer Review Reports

The CIE reviewers shall complete an independent peer review report in accordance with the requirements specified in this PWS and OMB guidelines. Each CIE reviewer shall complete the independent peer review according to required format and content as described in **Annex 1**. Each CIE reviewer shall complete the independent peer review addressing each ToR as described in **Annex 2**.

Other Tasks – Contribution to Summary Report

The CIE reviewers will assist the Chair of the panel review meeting with contributions to the Summary Report, based on the terms of reference of the review. The CIE reviewers are not required to reach a consensus, and shall provide a brief summary of each reviewer's views on the summary of findings and conclusions reached by the review panel in accordance with the ToRs.

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, passport number, country of passport, travel dates, country of citizenship, country of current residence, and home country) to the NMFS Project Contact for the purpose of their security clearance, and this information shall be submitted at least 30 days in accordance with the NOAA Deemed Export Technology Control Program NAO 207-12 regulations available at the Foreign National Guest website. The contractor is required to use all appropriate methods to safeguard Personally Identifiable Information (PII).

Place of Performance

The place of performance shall be at the contractor's facilities, and in Tutuila, American Samoa or virtually dependent on conditions of the COVID 19 pandemic.

Period of Performance

The period of performance shall be from the time of award through April 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 two weeks of award	Contractor selects and confirms reviewers
Approximately 2 weeks later	Contractor provides the pre-review documents to the reviewers
February 17-23, 2023	Each reviewer participates and conducts an independent peer review during the panel review meeting
Within two weeks of panel review meeting	Contractor receives draft reports

Within three weeks of receiving draft reports	Contractor submits final reports to the Government

*The Chair's Summary Report will not be submitted to, reviewed, or approved by the Contractor.

Modifications to the Performance Work Statement: 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. The PWS and ToRs shall not be changed once the peer review has begun.

Applicable Performance Standards

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

Travel

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

Restricted or Limited Use of Data

The contractors may be required to sign and adhere to a non-disclosure agreement.

NOAA Fisheries Project Contact:

Felipe Carvalho
NOAA Fisheries
FRMD/PIFSC/NMFS/NOAA
1845 Wasp Boulevard, Bldg. #176
Honolulu, Hawaii 96818
Felipe.Carvalho@noaa.gov

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

1. The report must be prefaced with an Executive Summary providing a concise summary of the findings and recommendations.

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.

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

4. 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.

5. Reviewers should elaborate on any points raised in the summary report that they believe might require further clarification.

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

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

8. The report shall include the following appendices:

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

Annex 2: Terms of Reference for the Peer Review

External Independent Peer Review under the Western Pacific Stock Assessment Review framework: 2023 Benchmark Stock Assessment for the American Samoa Bottomfish

For questions 1-9 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. If applied, is the choice of indicator species to evaluate more poorly known species that are in a stock complex appropriate?
10. 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 1-9 are "no"), indicate which alternative option should be used to inform setting stock status and fishery catch limits between 1) using the previous assessment, 2) using an indicator species, or 3) designing the stock status as "unknown".
11. As needed, suggest recommendations for future improvements and research priorities. Indicate whether each recommendation should be addressed in the short/immediate term (for this assessment), mid-term (next assessment) 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.

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

Annex 3: Tentative Agenda

2023 Benchmark Stock Assessment for the American Samoa Bottomfish

Tutuila, American Samoa

February 17-23, 2023

Day 1, Friday February 17

- Welcome and Introductions
- Background information
    - Objectives and Terms of Reference
    - Fishery operation
    - Management
- History of stock assessments and reviews
- Data
- DMWR creel-survey and Biosampling
- Public comment period (30 minutes)

Day 2, Saturday February 18

- Presentation and review of stock assessment

Day 3, Sunday February 19

- Continue review of stock assessment

Day 4, Monday February 20

- Continue review of stock assessment
- Public comment period (30 minutes)

Day 5, Tuesday February 21

- Continue review of stock assessment

Day 6, Wednesday February 22

- Continue review of stock assessment
- Panel discussion (Closed; afternoon)

Day 6, Thursday February 23

- Present results (Morning)
- Public comment period (30 minutes)
- Adjourn

Appendix 3. List of Participants

WPSAR Panel
Erik Franklin University of Hawaii
Joseph Powers CIE
Patrick Cordue CIE

NMFS - Pacific Islands Fisheries Science Center
Robert Ahrens
Felipe Carvalho
Marc Nadon
Megumi Oshima
Marlowe Sabater

NMFS - Pacific Islands Regional Office
Brett Schumacher

Territorial Agency
Taotasi Archie Soliai (DMWR)
Domingo Ochavillo (DMWR)
Warren Sevaaetasi (DMWR)
Christina Samau (DMWR)
Tepora Lavatai (DMWR)
Yvonne Mika (DMWR)
Selaina Vaitautolu (DMWR)
Mareko Milo (DMWR)
Letisha Fala (DMWR)
Shaun Laolagi (DMWR)
Auvaa Soonaolo (DMWR)
Herbie Umi (DMWR)
Tony Langkilde (DOC)

Western Pacific Fishery Management Council
Mark Daniel Fitchett
Nonu Tuisamoa (Advisory Panel)

American Samoa Fono (legislature)
Samuel Meleisea

American Samoa Fishing Community
Omar Shalhout
Keith Ahsoon
Leuma Sue
Fereti Lemoa
Punipuao Lemoa
Ropeti Misa
Howard Dunham
Calvin Ilaoa
Muamalae Tata Aga
Maselino Ioane
Ogesefolo Tuala
Paepae Simi
Manaima V
Brian Peck