Report on the

Independent peer review of length-based stock assessment methods for coral reef fish stocks in Hawaii and other U.S. Pacific territories

Prepared for:
The Center for Independent Experts

Dr Kevin Stokes
Stokes.net.nz Ltd
59 Jubilee Rd
Khandallah
Wellington 6035
New Zealand
Ph: +64 (04) 973 7305
E-mail: kevin@stokes.net.nz

EXECUTIVE SUMMARY

The CIE independent report shall be prefaced with an Executive Summary providing a concise summary of the findings and recommendations following Annex 2 Terms of Reference questions.

The review was of length-based stock assessment methods for coral reef fish stocks in Hawaii and other U.S. Pacific territories. The methods comprise multiple components, with many decisions to be made within each. The components are: A) use of length composition data and life history relationships to develop distributions of life history parameters to be used in YPR and SPR analyses; and B) combine outputs from (A) with estimates of abundance to calculate OFL/ABC. At component (A), decisions need to be made on: i) length composition data sources and judgments need to be made from those, notably on length at first capture; and ii) whether to use one of a three tier hierarchy of methods to derive life history distributions and thence YPR and SPR estimates. Those three, component (Aii) methods, in order of likely preference, draw on: i) raw, species specific life history data; ii) borrowed, valid life history estimates and CVs; or iii) data from multiple species and taxa to develop general, taxa level multivariate life history distributions (the "data poor stepwise" approach). At component (B), decisions need to be made regarding abundance estimation based on: i) visual survey data; or ii) commercial and recreational catch data. There are multiple sources of uncertainty at all stages of the complex approach.

Overall, the logic of the multi-step approach is sound and use of life history information and length composition data potentially provides a way of moving towards: i) status determination, and, possibly, though with some difficulty, ii) OFL/ABC. However, the method is not a standalone approach that can be automated or for which a simple guide can lead to reliable and replicable outputs with known biases and variances. It requires extensive input by and interaction with analysts, data providers, and, ideally, others. To the extent that a decision tree to guide usage and ensure transparent recording of decisions can be developed, this would be helpful. Ideally, the approach would be subject to simulation testing to understand the conditions under which it would provide valid management inputs; this, however, would be complex given the large number of component parts and decisions about data usage and it is not clear that it is warranted.

The data poor, stepwise approach (Aiii, above) is novel and has merit, but it is unclear how robust the currently-derived multivariate life history distributions are. It would be helpful to see in detail the basis of data selection and filtering and the statistical fitting used. This was not considered in detail during review.

In practice, the overall approach is fraught with problems and should only be used case-specifically and with great care. Decisions at every point need to be clear and justifiable, and appropriate sensitivity testing would be required to test the robustness of management–related outputs. At many points, it might be preferable to use group approaches to decision-making on data usage and parameterization. Given the complexity of the overall approach, group approaches can build more confidence in outputs for management purposes.

At this stage, while the approach is logical, it is not possible to say it is robust, either generally or, especially, case specifically.

BACKGROUND

The main body of the report shall consist of a ***Background****, Description of the Individual Reviewer's Role in the Review Activities, Findings of whether they accept or reject the work that they reviewed, and an explanation of their decisions (strengths, weaknesses of the analyses, etc.) for each ToR, and Conclusions and Recommendations in accordance with the ToRs.*

Management

As noted in multiple, provided references (e.g., Nadon and Ault, Submitted 1), coastal, tropical, reef fisheries are characterized by highly diverse catches, and often need to be managed based on limited human and financial resources. For the large majority of such fisheries, data are therefore limited, and it is necessary to make management decisions not just in the face of uncertainty but potentially even in the face of near ignorance. Many management systems, however, have evolved to make decisions based on increasingly complex technical advice based on increasing types and volumes of data. The United States is no exception; for Federally-managed fisheries, the 2006 Magnuson-Stevens Reauthorization Act (MSA) requires regional Fishery Management Councils (Councils) to specify a mechanism for specifying Annual Catch Limits (ACL). By default, all species listed in Fishery Management Plans (FMP), unless identified as an "*ecosystem component species*", are considered to be in fisheries and ACL are required to be set. This sets a high bar and difficult technical requirement.

ACL are required to be set by Councils, cognizant of scientific advice on Overfishing Limits (OFL) and Allowable Biological Catches (ABC), where the distance between OFL and ABC is determined allowing for scientific uncertainty. Councils can set ACL no higher than the ABC, and equal to the ABC only with clear justification. ACLs are used to trigger Accountability Measures (AM) aimed at ensuring ACL are not exceeded or, sometimes, that an Annual Catch Target (ACT) might be achieved. Clearly, the Federal system is heavily weighted towards catch-based (output) management, with a definitive need for certain types of metrics from scientific processes. For data-limited fisheries this is a problem, especially so when the primary data limitation is poor catch recording.

Councils are free to develop their own systems within the constraints of the MSA. The Western Pacific Fishery Management Council (WPFMC) has developed a five-tiered system, with Tiers 1 through 5 relating to progressively more poorly informed stocks. The Tiers determine how ABCs are to be set dependent on the reliability of and uncertainty in estimates of OFL.

This review is not intended to consider the general, Federal OFL-ABC-ACL system or the WPFMC Tier system. The foregoing, however, sets the tone for the review in that it provides context for application of the methods under consideration.

Hawaiian, coastal reef fisheries are not managed exclusively by the WPFMC. Indeed, the majority of commercial and recreational catch is taken within 2 nautical miles of the coast (except westward from Moloka'i, on Penguin Bank); commercial landings are monitored by the Hawaii Division of Aquatic Resources (HDAR); and State management of reef fish does not include catch-based measures. No background was provided on State measures prior to the review, but an outline of current measures was provided on request. Measures in effect (if not 'in force') include size limits for about 17 species, valid statewide, and minimum mesh sizes and bag limits for some species (including parrotfishes, *Scaridae*, and goatfishes, *Mullidae*). Marine Protected areas (MPA) have been established since 1967, covering about 0.4% of the Main Hawaiian Islands (MHI) coastline, while partial MPAs cover further 3.6% and restricted Access Areas cover 6.5%.

It is important to note that the Federal system requires scientific advice on stock status and, specifically, on OFL and ABC. The State system does not require advice on potential catches, but could in principle use outputs from methods under consideration to set input controls such as size limits, and/or seasonal or area restrictions in response to status determination. This review does not consider management, and the ToR do not refer explicitly to OFL. Nevertheless, given presentations made during the review, it is clearly necessary to consider the methods with respect to status determination utility and ability to set OFL/ABC.

The methods under review aim to use fundamental understanding of life-history theory to enable length-based measures to be converted in to growth and natural mortality estimates, and hence derive yield *per* recruit-based advice on stock status (e.g., %SPR, F/Fmsy), potentially size limits, and, if additional information is available on stock size, catch limits. The review is tasked with considering the methods and general applicability, as well as implementation in the Hawaiian context. It is specifically tasked with considering methods given acceptance of data for management purposes. That is, not considering the data provision and quality as part of the review. These data include those derived from biological sampling, visual dive surveys, recreational catch monitoring, and commercial catch monitoring. All of these data have uncertainty associated with them. Part of the review remit is to consider if uncertainty has been properly accounted for. While not reviewing the data sources *per se*, it is inevitable that some consideration has to be made of those sources.

Of most importance for setting OFL, are data on recreational and commercial catches, and/or dive surveys. Either is needed to enable absolute biomass estimation in order to set OFL, but neither is considered here in any detail. Regarding catch records, as noted e.g. by STAR (2011), considering catch-based data-poor methods but applicable to any use of catch records to determine OFL, "*catch-only methods rely on the availability of historical catches, and development of OFLs is nearly impossible without this information*"; the exception being if an estimate of absolute biomass is available from, e.g., a survey. There are known problems with both recreational and commercial catch statistics in Hawaii. Some difficulties with commercial records were traversed during WPSAR (2009) and it is not apparent that the problems identified at that time have been resolved or are irrelevant for the reef fisheries. Dealer licensing is yet to be mandated, records are of landings rather than catches, length sampling is poor, etc. The majority of reef fish catch is by recreational fishers and there are multiple issues with recreational catch estimation through MRIP.

The alternative to catch-based OFL estimation is to use surveys as a means of estimating absolute biomass. The majority of the visual survey data used for assessment and potential OFL estimation purposes comes from the Pacific Reef Assessment and Monitoring Program (Pacific RAMP), set up as a broad scale ecological status and trend monitoring program. The program operated initially on an occasional basis before becoming biannual and now triannual, with extensive coverage of US-associated coral reef islands and atolls in the Western Pacific Ocean. Additional dive survey coverage by Reef Fish Survey (RFS) cruises focused on populated islands, including the Main Hawaiian Islands (MHI), has been funded by NMFS in order to increase the dataset available for stock assessment purposes. The dive surveys use 7.5m radius 'cylinders' laid along a transect in depths to 30m on hard bottom areas, within which divers sequentially count fish by species using differing, species specific protocols. The design is random stratified. Estimation efficacy of biomass will vary by species and is subject to a wide range of uncertainties. Detailed consideration is beyond the scope of the ToR for this review.

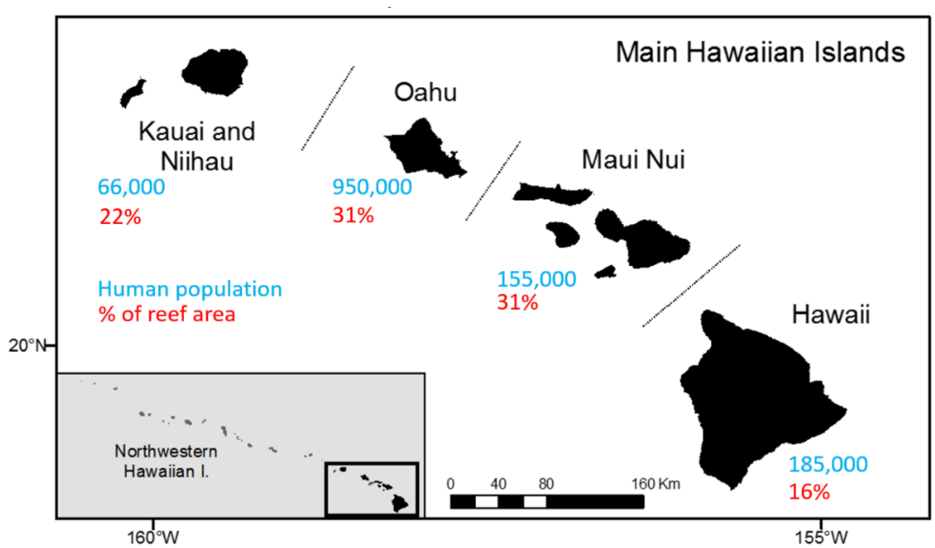
Stocks and Fisheries
The Hawaiian reef fishes caught in recreational and commercial fisheries comprise many hundreds of species, across multiple taxa. The key taxa caught are *Acanthuridae* (surgeon or unicornfishes), *Carangidae* (pompano/jacks), *Mullidae* (goatfishes), and *Scaridae* (parrotfishes), *Holocentridae* (squirrelfishes), as well as *Lutjanidae* (snappers), and *Lethrinade* (Emperors).

Recreational catches (as outlined in the review) account for the majority of the statewide catch, but are caught more diffusely than commercial catches. The table below shows ratios of recreational to commercial catches for the period 2004-2009, based on estimates from the NOAA Marine Recreational Information Program (MRIP) and commercial catch reports processed by HDAR.

Family	2004-2009 annual mean catch (lb)		Ratio
	Commercial Catch	MRIP Re-assessment	
Acanthuridae	86,463	245,333	2.8
Scaridae	40,450	82,075	2.0
Mullidae	46,880	216,472	4.6
Holocentridae	43,436	18,856	0.4
Kyphosidae	28,228	61,068	2.2

Recreational fishing effort is variable by area. No recreational fishing takes place in the NW Hawaiian islands (NWHI), and according to information provided in the review, effort is split amongst the Main Hawaiian Islands (MHI) as: Oahu, 47%; Hawaii, 28%; Maui and Nui, 17%; and Kauai and Niihau, 9%.

The following graphic, taken from presentation during review, shows the MHI island groups with population and estimates of percentage reef area. It is clear that the effort to reef area ratios between the islands are variable, with higher fishing intensity *per* reef area in Oahu and Hawaii than the other islands.



REVIEW PROCESS

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

The review organization, communication, distribution of documents for reading, etc., were all straightforward. The venue for the review was uncertain in advance and during the meetings, with changes to rooms made daily. This did not, however, pose any major difficulties given the small number of participants (see Appendix 3) and good on-the-spot communication. The general venue location (the University of Hawaii at Manoa) was adequate but not ideal, with basic facilities, no internet provision, and an often noisy environment – due both to excessive background noise from student dining facilities and from ageing air conditioning or fans.

The review was hosted by the NMFS PIFSC. I have no reason to think the review was not adequately publicized, but note participation was very limited (see Appendix 3). On Day 1, in addition to the PIFSC primary analyst (Nadon) and three CIE reviewers (Dichmont, Pilling, and Stokes), 7 PIFSC staff and one NOAA PIRO staff member were present. No WPFMC or State fishery management representatives participated and there was no public participation. On following days, the participation was further reduced. Only on Day 2 was there a WPFMC presence, and at no point through the meeting was there any State or public participation. Daily participation is indicated in Appendix 3. One area of difficulty for the review was to put it in to a full management context to allow a clear interpretation of the ToR (see below). WPFMC and State participation would have been helpful in this respect.

No Chair was appointed. On Day 1, PIFSC (Brodziak) welcomed participants, gave administrative information, and provided an initial overview and comment on the Terms of Reference (ToR). Thereafter, the review worked informally. This did not pose any problems, with all presenters and reviewers contributing in a friendly, relaxed and transparent manner. Discussion was wide-ranging where useful but also well-focused as necessary. The primary presenter (Nadon), whose work was the key subject of review, acted professionally and openly throughout and should be commended for his excellent attitude to the review. Few requests were made by reviewers but all were completed efficiently and with clear and timely responses. Only two additional PIFSC staff were present for all of Days 2 through 4 of the review (Yau and Boggs); both provided useful and open input and their continued presence was appreciated, providing helpful contributions as well as collegial support for Nadon.

While limited in participation, I appreciate the openness and friendliness of the PIFSC staff and the well-natured feel of the review. I am grateful to PIFSC staff and to fellow reviewers for what I consider to have been a thorough but enjoyable review.

Terms of reference, distributed in advance as part of the Statement of Work (see Appendix 2), were, in my opinion, clumsy. The panel has interpreted ToR 1 and 2, and ToR 3 and 4 as pairs, with the first in each pair referring to general methods, and the second in each pair to Hawaiian implementation. There is not a clear distinction between the ToR in each pair and this is necessarily reflected in responses to ToR, below.

For ToR 3 and 4, related to management, it would have been very helpful to have clearer input from both Federal and State management agencies, outlining decision-support needs and consequent review expectations. Only a general outline of the WPFMC approach to setting of OFL, ABC, ACL, AM and the Tier system was provided as a presentation on Day 1. No management background materials were provided, nor were any management documents provided during the review. ToR 3 refers explicitly to the FEP (Fisheries Ecosystem Plan), but this was not made available in advance or even during the meeting. A clearer description of Federal and State management of the reef fisheries would have been useful. I note also that ToRs 3 and 4 refer specifically to status metrics ("*estimated population benchmarks and management parameters (e.g., spawning potential ratio, F/Fmsy, B/Bmsy, stock status)")* rather than to OFL, ABC, etc. It is not clear if "*management goals stated in the FEP*" is intended to cover OFL, etc., but such language would not normally be interpreted in this way, with goals being higher level than setting of e.g. OFL. As the review included substantial material on using methods to estimate OFL, greater background and clarity as to ToR intent would have been useful.

Papers distributed in advance covered fundamental aspects of the data-poor methodology to be considered. However, like the ToR, none of those papers included any indication that the review would extend to consideration of OFL-setting. Also, the papers provided did not include, nor had any reference to, other key papers such as Kritzner *et al.* (2009).

Regarding ToR6, it was made clear at the outset of the review meeting that there would be no WSPAR Panel Report required. However, the panel agreed that on the final day, a verbal report on conclusions and findings, addressing each ToR, would be given. This was done (see below).
I note that no specific time was given during the review to consolidate thoughts on ToR5 and no specific feedback was given on recommendations during the verbal report under ToR6. ToR 5 and 6 would likely have been better attended to if the meeting had been formally chaired, but overall my sense of the meeting is that this was not problematic.

I recommend that future WPSAR ToR are crafted to provide more clarity and that, in particular, ToR are explicit as to what management-related assessment outputs are required to be reviewed. Similarly, documents provided in advance need to address the ToR in all respects.

REVIEWER'S ROLE IN THE REVIEW ACTIVITIES

*The main body of the report shall consist of a Background, Description of the **Individual Reviewer's Role in the Review Activities**, Findings of whether they accept or reject the work that they reviewed, and an explanation of their decisions (strengths, weaknesses of the analyses, etc.) for each ToR, and Conclusions and Recommendations in accordance with the ToRs.*

As required by the CIE Statement of Work (see Appendix 2), I read all background materials provided (see Appendix 1), participated fully during the panel review meetings in Honolulu, and have provided an independent report (this document). During the review, I worked with fellow CIE reviewers during the meetings and also during extensive evening discussions. As a panel, we agreed (unanimously) feedback to be provided on the final day. Pilling (CIE) made notes of our agreements as a basis for presentation, modified during the feedback session. Dichmont (CIE), led the verbal feedback, based on the notes, with Pilling and myself contributing fully.

FINDINGS BY STOCK AND ToR

The main body of the reviewer report shall consist of a Background, Description of the Individual Reviewer's Role in the Review Activities, ***Summary of Findings for each ToR in which the weaknesses and strengths are described****, and Conclusions and Recommendations in accordance with the ToRs.*

1. Review the assessment methods used: determine if they are reliable, properly applied, and adequate and appropriate for the species, fisheries, and available data considering that the data itself have been accepted for management purposes.

AND

2. Evaluate the implementation of the assessment methods: determine if data in its current form are properly used, if choice of input parameters seems reasonable, if models are appropriately specified and configured, assumptions are reasonably satisfied, and primary sources of uncertainty accounted for.

The methods under consideration relate to: a) estimation of growth and natural mortality from length-based and life history information sources; b) use of generated estimates of life history parameters to estimate status metrics (%SPR, F/Fmsy) using population simulation; and c) abundance estimation in order to estimate e.g., OFL and ABC.

Assessment methods papers provided in advance are Nadon and Ault (Submitted 1,2) and Nadon *et al.* (2015), as well as relevant background papers (see Appendix 1). Nadon and Ault (Submitted 1) describes and tests a stepwise, stochastic simulation approach to estimating missing life history parameters by building multivariate probability distributions of life history distributions based on life history estimates (not raw data) from studies on multiple species across six taxa. Testing is through comparison of estimated probability distributions of life history parameters, and generated SPR distributions for three species to distributions generated by bootstrapping from raw data for those species.

Nadon and Ault (Submitted 2) further implements methods to estimate SPR for a range of Hawaiian reef fish stocks by using length composition data from visual dive surveys and the methods in Nadon and Ault (Submitted 1). Distributions of life history parameters and SPR are calculated for 8 species for which length composition data are available, but no life history information, and a further 17 using the data poor approach but for which it is possible to compare life history distribution estimates with previous assessments.

Nadon *et al.* (2015) attempts status determination for 19 species for which raw data or borrowed data are available (but does not use the data poor methods described in Nadon and Ault (Submitted 1)).

No papers provided in advance considered OFL/ABC estimation. The review considered OFL/ABC calculation using population simulation outputs in association with abundance data derived from visual dive surveys and/or catch records. The review is not intended to cover the abundance estimation approaches *per se* and the issues arising (esp. accounting for uncertainty) are dealt with at ToR 3 and 4.

During the review, presentations relevant to ToR 1 covered the generation of life history parameters and distributions from: i) local studies with raw data; ii) estimates and CVs borrowed from other areas, species, etc.; and iii) the stepwise "data poor" approach described in Nadon and Ault (Submitted 1) and used for 17+8 species in Nadon and Ault (Submitted 2). Within these methods, life history-based or maximum age-based estimation of natural mortality, and alternative data sources of size-composition data (dive surveys and commercial catch samples) were also considered. These issues span ToR 1 and 2.

Use of length-based approaches to derive estimates of status against benchmarks have been described and used in other regions. The review included relevant papers from other studies, as fully referenced in Nadon and Ault (Submitted 1, 2) and Nadon *et al.* (2015). Those various studies have explored sensitivities and dependencies, and there is good reason to accept such methods as a way of estimating benchmarks, if carefully applied. Nadon and Ault (Submitted 1,2) go further than the published studies in attempting to use the life-history approach in the absence of local data (i, above) or even specific alternative studies from which to borrow life history estimates (ii, above).

While not explicit in the various documents provided, but clear enough from presentations, there is a hierarchy of approaches that should be followed to derive life history estimates and status estimates for management purposes. If raw data are available for a species (i, above), use the approach used in referenced studies, and as used for testing the data poor approach in Nadon and Ault (Submitted 1) or for estimating life history and SPR distributions in Nadon and Ault (Submitted 2) and Nadon *et al.* (2015). Absent local data, if there is reason to accept specific surrogate data or life history estimates from elsewhere, this would be the next preferred option (ii, above). As noted in presentations, borrowing life history estimates usually comes without accompanying raw data from which to bootstrap, and use of other information on error distributions is necessary. As shown in presentations, the work of Kritzner *et al.* (2001), based on four reef fish species from the Great Barrier Reef (GBR), might be useful in this respect. However, great care is needed in making inferences either from particular stocks or about coefficients of variation derived from the GBR stocks; none may be suitable. The Krtizner *et al.* study is for four species only, and it may not be useful in many cases. Use of borrowed data may therefore be compromised, leaving only the data poor, step-wise approach available (iii, above). The Kritzner *et al.* study is not referenced in any of the primary review documents, and it is unclear how this was handled in the Monte Carlo simulations.

Only if local life history information or relevant information from other stocks were not available, would it be logical to move to use the data poor method described in Nadon and Ault (Submitted 1). The question, if such use is made, is does the method provide a reliable basis for status determination and, potentially, OFL/ABC estimation?

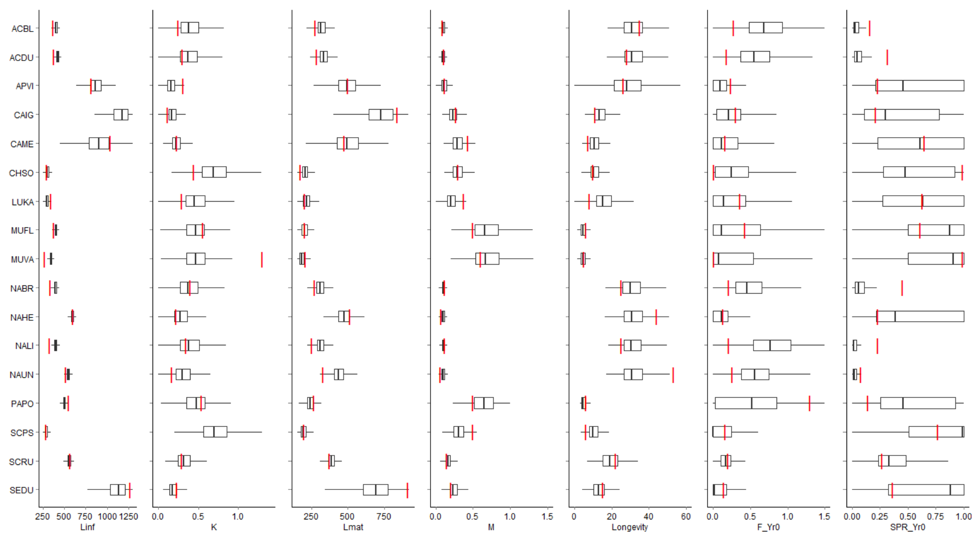
Normally, when considering novel methods, it is possible to work from first principles or simulation studies to determine properties of a method and advise on applicability (see, e.g., Wetzel and Punt, 2011 for a data poor application). Detailed implementation choices based on automated technical decisions can then be tested and specified. In the case at hand, however, this is not the case. Nadon and Ault (Submitted 1) describes multiple choices to be made *inter alia* on meta data selection (taxa, species, specific studies, locality, environment, male *vs* female maturity, means of estimating natural mortality (*via* maximum observed length and allometric relationships or longevity), and a number of other subjective issues). Once data were selected, numerous decisions were also made when fitting multivariate statistical relationships between life-history parameters, with ultimate selection based on "*in depth,*" but nevertheless expert/subjective examination of scatter plots. Natural mortality is estimated for all stocks using a relationship reported in Nadon *et al*. (2015) in which M (assumed equal to Z for the unexploited NWHI) is regressed on the inverse of published values of maximum age to obtain a survival coefficient. That study suggests a survival constant of 0.043, and Nadon and Ault (Submitted 1,2) use a value of 0.05 for their data poor, stepwise analysis. The survival value is large compared to the 0.015 due to various studies by Hoenig and others. Once the data selection and fitting steps were made, Nadon and Ault used their accepted relationships for six separately fitted taxa in a "data poor" approach to generate probability distributions of species specific life history parameters based on input values of Lmax, as the 99th percentile of lengths observed in the RAMP/RFS surveys (another choice). Results obtained using the data poor approach were compared to those obtained for the same three species with life history parameters bootstrapped from raw data for the three species (the "data rich" approach). Comparisons between the data rich and data poor approaches were made for estimated probability distributions of life history parameters, and also for derived %SPR estimates assuming length at first capture estimated subjectively for each species from data, and the natural mortality estimates using the survival constant of 0.05. Only three species were used for comparisons, although multivariate life history distributions were developed for six taxa and could be developed for more. It is therefore difficult to draw conclusions about general methodological efficacy.

The procedures outlined by Nadon and Ault are logical and reasonable, but it is difficult to assess the approach as a consolidated stock assessment method. This is because results are dependent on so many decisions in data selection, initial statistical development, and then implementation, and are available for just three species. The test results show the expected and appropriate feature of increasing uncertainty when the data poor approach is used, but do not provide a basis for systematic analysis of sources of bias and uncertainty. Species specific variations in estimated probability distributions between data rich and data poor approaches may be because of poor metadata availability/selection, choices made when fitting statistical relationships by taxa, method of estimating natural mortality, choice of length at first capture, etc.

During review, care was taken to explain choices related to natural mortality estimation, use of the 99th percentile of lengths from RAMP surveys, etc., but there was no consideration of the statistical fitting of life history estimate meta data used to generate the multivariate probability distributions that serve as the base for application of the data poor approach. Nadon and Ault (Submitted 1) describes well the general considerations of the process, but without careful scrutiny it is not possible to say if the base is sound, nor if it fully captures uncertainties that will flow through to derived estimates.

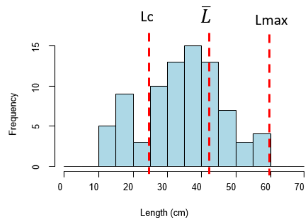
Results for the 17 species reported in Nadon and Ault (Submitted 2), and expanded during review, provide a slightly fuller picture of possible reliability. The graphic below was produced on request, during review; it is an extended version of Figure 4 in Nadon and Ault (Submitted 2). It shows estimated distributions for multiple life history parameters (as [5, 25, 50, 75, 95] percentile box and whisker plots) using the data poor approach, with superimposed red bars showing median values from previous growth studies (and implied SPR). Without attempting detailed analysis by taxa, etc., it is clear from the final column that %SPR distributions generated using the data poor approach may be biased in either direction and may be very wide.

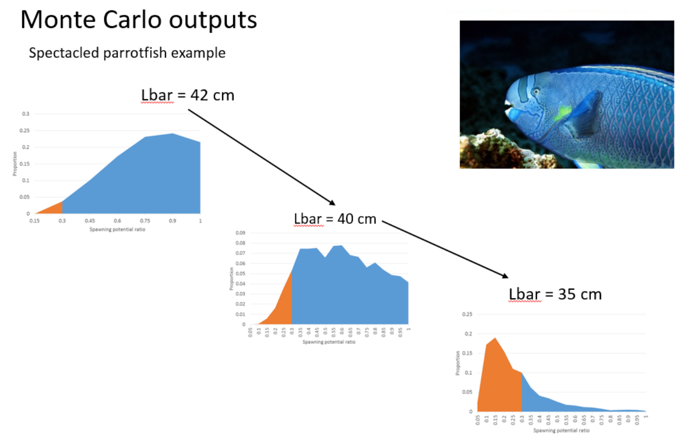
It is difficult to look in detail, but I note some species for which results may be at odds with those shown elsewhere. For example, for *Aprion virescens* (APVI, third row down), the SPR shown in the final column has a median circa 0.5. In Nadon *et al.* (2015), the SPR in Table 3 and Figure 7 is 23%, and in presentation, there seem to be discrepancies in some of the life history parameter distributions for APVI. This is unclear.



Ideally, the approach would be tested systematically using developed data sets with known properties (e.g., Wetzel and Punt, *ibid*). Such work would be difficult and extensive, and would require automation of many of the decisions taken and assumptions made by Nadon and Ault. I am reluctant to recommend such an analysis, but note it would have wider utility than for the Western Pacific region. Nadon and Ault (Submitted 1) provides thorough text on the processes used to underpin the data poor, stepwise method development and implementation, but it is not possible to definitively say the method *per se* is *"reliable, properly applied, and adequate and appropriate for the species, fisheries, and available data"*. I strongly suspect that it is carefully conducted work, and the review discussions showed clearly that the work had been undertaken with great care and attention to detail. The primary analyst (Nadon) is very clearly fully "on top" of the work conceptually and in detailed application. In the absence of raw, species specific data, use of the approach has great appeal as it provides, in principle, a means of moving beyond catch-only data poor methods (for which other assumptions have to be made to calculate status metrics and OFL/ABC), and of utilizing a direct and intuitively interpretable measure (mean length) as an indicator of stock status and fishing pressure.

Areas of judgment (subjectivity) need special care. For example, the choice of length at first capture is not automated but is an expert call, case by case. During review, an example was given for the parrotfish *Scarus perspicillatus*, showing selected length at first capture, maximum length, and consequential mean length of the selected length range (see below, left hand panel). Maximum length is the starting point for the stepwise generation of species specific life history distributions based on the multivariate analysis outputs. SPR estimation depends on those outputs as well as on the input length at first capture. The distribution of SPR is highly dependent on length at first capture, as shown during review and below. Choice of length at first capture may not be clear in the data and the size of data bins will also be a determinant of choice of first, and hence, mean length.





In my view, the data poor approach should be applied case specifically and with care, with clear exposition of decisions made at all points in the analysis for each species. Though data poor, this would entail substantial effort, especially for application to multiple species. Stock assessment methods ideally can be applied automatically or at least with clearly guided and recorded technical decisions at key points. As a method, the data poor, stepwise approach does not yet fit that bill, requiring multiple decisions, many subjective or expert, all with consequential sensitivities. It is important that any method should be replicable and robust to changes in analysis, etc. As currently put forward, the stepwise approach is not yet so well defined.

I note SPR methods are standard, and see no need to comment on these. However, dependency on estimates of M, length at knife-edge selection, and other parameters are all critical in determining SPR and need to be carefully considered.

The methods used to estimate biomass as a basis of OFL/ABC calculation are also outside the scope of this review. As noted in the background, there are known problems with both recreational and commercial catch data in Hawaii, and visual dive surveys may or may not, species by species, provide a good basis for determining absolute biomass. The surveys are not designed specifically for absolute abundance estimation. Total habitat surveyed is a very small percentage of all hard ground within the 30m contour; habitat suitability is considered constant; some species are distributed to much greater depths than 30m; species have variable crypticity, size, speed; some species will have simple, temporally invariant distributions, while others will be highly variable spatially and temporally; etc. A thorough, species by species consideration would be needed to tease out for which species the surveys might provide consistent and unbiased estimators.

3. Comment on the scientific soundness of the estimated population benchmarks and management parameters (e.g., spawning potential ratio, F/Fmsy, B/Bmsy, stock status) and their potential efficacy in addressing the management goals stated in the relevant FEP or other documents provided to the review panel.

AND

4. Determine whether the results (such as SPR-based reference points, stock status) in their current form from the assessment methods can be used for management purposes without further analyses or changes considering that the data itself [sic] have been accepted for management purposes.

It was not made clear in the review how to interpret the ToR with respect to "*management goals stated in the relevant FEP".* However, presentations during review, though not documents provided in advance, included estimation of OFL and ABC, and the panel was advised the words in ToR 3 should be interpreted in that light. As noted previously, the FEP was not provided.

YPR and SPR methods are standard, and I see no need to comment on these. They are appropriate so long as the input parameters are, and are commonly used to determine stock status in the US. Definitions and benchmarks (esp. 30%SPR, with 50% probability) are normalized. The approach taken to OFL/ABC calculation is to bring together the YPR estimates with estimates of absolute abundance from visual dive surveys or from commercial and recreational catch. During review, results based on both abundance estimates were conflated. In principle, I see no problem with either approach separately, though do not think it appropriate to bring the two estimates together. If a fishery independent absolute biomass estimate is reliable, it makes sense to use it. If only catch estimates are available, then, in principle, the length-based methods used to determine status provide more information to the OFL-setting process than catch-only methods, which are becoming widespread, but which require major assumptions about depletion to be made (e.g., McCall, 2009; Dick and McCall, 2011), as well as requiring other inputs (e.g., natural mortality) which are also required for use when deploying length-based methods.

Following the logic at ToR 1 and 2, a hierarchy of use is warranted. For the length-mortality modelling that feeds into simulation to estimate Fmsy and SPR, decisions need to be made on: a) use of life history 'data', and b) which length composition data to use to estimate length at first capture and mean length of (the fishery selected) population. A decision tree is needed with as much automation, and then guidance as to judgments as may be possible. All decisions need to be transparent and replicable. At (a), as noted at ToR 1 and 2, the hierarchy should be: i) use the raw data from local studies where available; ii) use surrogate data on life histories (and CVs) where available and if appropriate; and iii) use the data poor, stepwise approach (though see caveats at ToR 1 and 2). At (b), length composition data may be available from dive surveys or from commercial catch sampling. Both have strengths and weaknesses, generally and species by species, and preference will be species specific. If the hierarchy is followed and decisions are carefully made and transparent, then status metrics might be sound. Rigorous sensitivity testing to assumptions, especially subjective calls is necessary.

In practice, care is needed in detail, species by species, in order to determine if estimated metrics can be used for management purposes without further analyses. During review, no complete list of species was provided for consideration at ToR 4. What would be needed for each species under consideration is clarity as to decisions at (a) and (b), above, derivation of M, decisions on chosen length at first capture, maximum length choice, etc. For each species, these and relevant sensitivity tests are required before the metrics can be used.

The methods used to estimate biomass as an input to OFL/ABC calculation are specifically outside the remit of this review. As noted in the background, there are known problems with both recreational and commercial catch data in Hawaii, and visual dive surveys may or may not, species by species, provide a good basis for determining absolute biomass. The surveys are not designed specifically for absolute abundance estimation but rather for ecosystem monitoring. Total habitat surveyed is a very small percentage of all hard ground within the 30m contour; habitat suitability is considered constant; some species are distributed to much greater depths than 30m; species have variable crypticity, size, speed; some species will have simple, temporally invariant distributions while others will be highly variable spatially and temporally – the random stratification will therefore have different bias-variance trade-offs for each species; accuracy of visual length estimates with distance; etc. A thorough, species by species consideration would be needed to tease out for which species the surveys might provide consistent and unbiased estimators and to estimate suitable error distributions.

During review, a mean abundance estimate derived from the visual surveys was shown for the snapper *Aprion virescens*, with a CV of 22%. Intuitively, this is low. In order to use the survey-derived abundance estimates and CVs, I would recommend a species by species consideration of all factors relevant to estimating the uncertainty and reliability of the estimate. Methods used elsewhere to develop survey priors might be useful in this respect.

During the review meeting, partially worked examples showing estimated life history and OFL distributions were provided for only two species: the snapper *Aprion virescens*, and the goatfish *Parupeneus porphyreus*. None were provided in advance.

For *A. virescens*, a mean length distribution derived from commercial data is shown and used. No explanation of why commercial data were used rather than survey data is given. Neither is there any indication of length at first capture. Life history parameter distributions are shown, derived using the data poor approach and using the "borrowed data" approach, with data on life history taken from a 1980 study in New Caledonia, and using CVs from Kritzner *et al.* (2001). It is not clear if the New Caledonian study is relevant, and I cannot work from Kritzner *et al.* to the numbers shown (but trust the analyst). F and Fmsy distributions derived using both life history methods are provided. It is good that they are kept separate, but no indication is given for preference. Population abundance is then shown based on diver surveys, and from catch using both the data poor and borrowed data approaches. No indication is given of validity of the surveys or of catch reliability. Distributions of OFL are then shown for each life history method combined with catch or survey-derived abundance estimates. Finally, cumulative distributions of OFL are shown for both catch and survey cases, but with the life history-derived parameters from the data poor and borrowed data cases combined using a method due to Hamel (2015). I see nothing wrong with the sequence of OFL building, but find it unclear how decisions are or would be made case by case to use different component parts. There is a clear difference in OFL outcome depending on whether survey or catches are used, not just in the median (57 mt *vs* 105 mt), but also in the steepness of the cdf, which would affect ABC calculation. I do not think it appropriate to combine distributions from the data poor and borrowed data approaches. Either the borrowed data are appropriate or not. If they are, they should be preferred. If not, the data poor should be preferred (though it would not necessarily be valid).

As stated above, I think it is essential to provide a clear decision tree to guide usage of the general approach being advocated (i.e., use length composition data with life history information to provide YPR and SPR estimates of status, and then use those combined with abundance estimates to develop OFL/ABC). Then for each species, that tree needs to be transparent, with clarity as to why decisions are made and how sensitivities would arise. From experience at review, I would have considerable confidence in the primary analyst (Nadon), but that cannot translate into a simple confidence in the approach or results as might currently exist (but I am unclear as to what materials might, e.g., be sent for consideration by the WPFMC SSC).

5. Suggest research priorities to improve our understanding of essential population and fishery dynamics necessary to formulate best management practices. Comment on alternative data sources and modeling.

Management practice is jurisdiction dependent and what is 'best' can only exist in context. For the reef fisheries in question, commercial and recreational catch records are known to be poor, creating a major problem for any assessment method and, indeed, for management, which is largely dependent on catch monitoring in real time. Improvements in catch recording are to be encouraged, and might help in the future setting of OFL/ABC and meaningful ACL, but it is difficult to see how existing catch records can be used successfully in the overall assessment approach under review.

The dive surveys potentially provide a fishery independent way of estimating abundance that might be used in OFL/ABC-setting. However, if catches are not recorded in a timely and accurate fashion, it is hard to see the practical utility of OFL/ABC. Notwithstanding, if surveys are to be used for estimating absolute abundance, priority needs to be given to understanding whether or not this might be feasible, species by species, and, for each species, how the multiple sources of uncertainty need to be taken into account. The review did not consider this in detail, but to have confidence in estimates and associated uncertainty it is necessary if OFL/ABC are to be set.

Length composition data from surveys and, ideally, fisheries are needed to drive the length- and life-history-based data poor approach. Ensuring data collection should be a continuing priority.

On the life history components of the approach in question, it is clearly preferable to have local data. For key species (e.g., target, long-lived, threatened), collection of local life history data and life history studies should be a priority. A clear design of data collection and life history estimation for selected species by taxa could also provide a basis for local "borrowed" data or data poor approaches.

The data poor, stepwise approach is logically sound, but it is unclear how sensitive it is to data selection and assumptions made, or to the detailed statistical modelling. If the method is to be widely used to determine stock status of multiple species, or for OFL/ABC-setting, then it is important to potentially extend the work to include other taxa and more datasets. Also, before use, the 'innards' need to be looked at critically. The review did not do this (NB, I anticipated it would) and it should be a priority.

6. Draft a report of the WPSAR Panel conclusions and findings, addressing each Term of Reference.

It was made clear at the outset of the review meeting that there would be no WSPAR Panel Report. However, the panel agreed that on the final day, a verbal report on conclusions and findings, addressing each ToR, would be given. This was done; see sections above on Review Process and Reviewer's Role.

Conclusions and Recommendations

The main body of the report shall consist of a Background, Description of the Individual Reviewer's Role in the Review Activities, Findings of whether they accept or reject the work that they reviewed, and an *explanation of their decisions (strengths, weaknesses of the analyses, etc.) for each ToR, and* ***Conclusions and Recommendations*** *in accordance with the ToRs.*

Conclusions

The review was of length-based stock assessment methods for coral reef fish stocks in Hawaii and other U.S. Pacific territories. The methods comprise multiple components, with many decisions to be made within each. The components are: A) use of length composition data and life history relationships to develop distributions of life history parameters to be used in YPR and SPR analyses; and B) combine outputs from (A) with estimates of abundance to calculate OFL/ABC.

At component (A), decisions need to be made on: i) length composition data sources and judgments need to be made from those, notably on length at first capture; and ii) whether to use one of a three tier hierarchy of methods to derive YPR and SPR estimates. Those three, component (Aii) methods, in order of likely preference, draw on: i) raw, species specific life history data; ii) borrowed, valid life history estimates and CVs; or iii) data from multiple species and taxa to develop general, taxa level multivariate life history distributions (the "data poor stepwise" approach).

At component (B), decisions need to be made regarding abundance estimation based on: i) visual survey data; or ii) commercial and recreational catch data.

There are multiple sources of error at all stages of the complex approach.

Overall, the logic of the approach is sound and use of life history information and length composition data potentially provides a way of moving towards status determination, and even OFL/ABC estimation which is better informed than other data poor approaches. However, the method is not a standalone approach that can be automated or for which a simple guide can lead to reliable outputs with known biases and variances. It requires extensive input by and interaction with analysts, data providers, and ideally others. To the extent that a decision tree to guide usage and ensure transparent recording of decisions can be developed, this would be helpful. Ideally, the approach would be subject to simulation testing to understand the conditions under which it would provide valid management inputs; this, however, would be complex given the large number of component parts and decisions about data usage, and it is not clear that it is warranted.

The data poor, stepwise approach itself has merit, but it is unclear how robust the currently-provided multivariate life history distributions are. It would be helpful to see in detail the basis of data selection and filtering, and the statistical fitting. This was not done during the review.

In practice, the overall approach is fraught with problems, and should only be used case specifically and with great care. Decisions at every point need to be clear and justifiable, and appropriate sensitivity testing would be required to test the robustness of management–related outputs. At many points it might be preferable to use group approaches to decision-making on data usage and parameterization. Given the complexity of the overall approach, group approaches can build more confidence in outputs for management purposes.

At this stage, while the approach is logical, it is not possible to say it is robust, either generally or, especially, case specifically.

Recommendations

I **recommend** that future WPSAR ToR are crafted to provide more clarity and that, in particular, ToR are explicit as to what management-related assessment outputs are required to be reviewed. Similarly, documents provided in advance need to address the ToR in all respects.

To the extent possible, I **recommend** the development of a decision tree to guide usage and ensure transparent recording of decisions in application of the overall approach. The tree needs to apply to all components (*inter alia*, which length composition data; how length at knife-edge selection is chosen; which life-history distribution generator, and why; M; abundance scalar validity and choice; etc.).

I **recommend** that in practical application, the approach be reported in detail case specifically, rather than be applied in bulk. Decisions at every point need to be clear and justifiable, and appropriate sensitivity testing would be required to test the robustness of management–related outputs for each stock.

I **recommend**, in line with the preceding, that assessments would be more robust and credible if conducted using a group approach to decision-making on data usage and parameterization. Given the complexity of the overall approach, group approaches can build more confidence in outputs for management purposes.

I **recommend** further work on the data poor stepwise approach, to: i) extend the work to include other taxa, and ii) more datasets. Most importantly, ideally as part of an expansion, but otherwise on the existing work, the details of data selection and decisions, and statistical fitting, need to be reviewed before application for management purposes.

APPENDIX 1

BIBLIOGRAPHY

Prior to the Workshop, documents for reading were provided *via* e-mail. During the workshop, further papers were made available on request. A complete list is given below. Presentations were provided on the first morning of the review, and following requests made at the end of day one.

Papers circulated in advance:

- Ault, J.S., J.A. Bohnsack, and G.A. Meester (1998) A retrospective (1979-1996) multispecies assessment of coral reef fish stocks in the Florida Keys. Fishery Bulletin 96(3):935-414.
- Ault, J.S., S.G. Smith, and J.A. Bohnsack (2005) Evaluation of average length as an estimator of exploitation status for the Florida coral-reef fish community. ICES Journal of Marine Science 62:417-423.
- Ault, J.S., S.G. Smith, J. Luo, M.E. Monaco, and R.S. Appeldoorn (2008) Length-based assessment of sustainability benchmarks for coral reef fishes in Puerto Rico. Environmental Conservation 35: 221-231.
- N.M. Ehrhardt and J.S.Ault (1992) analysis of Two Length-Based Mortality Models Applied to Bounded Catch Length Frequencies. Transactions of the American Fisheries Society 121: 115-122.
- T. Gedamke and J.M. Hoenig (2006) Estimating Mortality from mean Length Data in Nonequilibrium Situations, with Application to the Assessment of Goosefish. Transactions of the American Fisheries Society 135: 476-487.
- Nadon, M.O, J.S. Ault, I.D. Williams, S.G. Smith, and G.T. DiNardo (2015) Length-Based Assessment of Coral Reef Fish Populations in the Main and Northwestern Hawaiian Islands. PLoS ONE 10(8): e0133960. doi:10.1371/journal.pone.0133960.
- Nadon, M.O. and J.S. Ault (Submitted 1) A stepwise stochastic simulation approach to estimate missing life history parameters for data-poor fisheries.
- Nadon, M.O. and J.S. Ault (Submitted 2) Assessment of data-poor Hawaiian coral reef fish populations using life history parameters obtained through a stepwise stochastic simulation approach.

Papers made available on request:

- E.H. Ellington, G. Bastille-Rousseau, C. Austin, K.N. Landolt, B.A. Pond, E.E. rees, N. Robar, and D.L. Murray (2015) Using multiple imputation to estimate missing data in meta-regression. Methods in Ecology and Evolution 2015, 6, 153–163.
- Kritzner, J.P., C.R. Davies,,and B.D. Mapstone (2001) Characterizing fish populations: effects of sample size and population structure on the precision of demographic parameter estimates. Can. J. Fish. Aquat. Sci. 58: 1557–1568.

- T.E. Raghunathan, J.M. Lepkowski, J.Van Hoewyk, and P. Solenberger (2001) A Multivariate Technique for Multiply Imputing Missing Values Using a Sequence of Re ression Models. Survey Methodology 27(1): 85-95.
- S. van Buren, and K. Groothuis-Oudshoorn (2011) mice: Multivariate Imputation by Chained Equations in R. Journal of Statistical Software 45(3): XXX.

Further references:

Francis, R.C.C.C. (2011) Data weighting in statistical fisheries stock assessment models. *Can. J. Fish. Aquat. Sci.* 68:1124-1138.

- Dick, E.J. and A, D McCall (2011) Depletion-Based Stock Reduction Analysis: A catch-based method for determining sustainable yields for data-poor fish stocks. Fisheries Research 110: 331–341.
- Hamel, O.S. (2015) A Method for calculating a meta-analytical prior for the natural mortality rate using multiple life history correlates. ICES J. Mar. sci. 72:62-69.
- MacCall, A.D. (2009) Depletion-corrected average catch: a simple formula for estimating sustainable yields in data-poor situations. ICES J. Mar. Sci. 66, 2267–2271.
- STAR (2011) Assessment Methods for Data-Poor Stocks. Report of Review Panel Meeting, Santa Cruz, 2011.
- Wetzel, C.R. and A.E. Punt (2011) Model performance for the determination of appropriate harvest levels in the case of data-poor stocks. Fisheries Research 110:342-355.
- WPSAR (2009) Report of the Western Pacific Stock Assessment Review 1: Hawaii DeepSlope Bottomfish, Honolulu, 2009.

Presentations available on Day 1:

- 0_Fishery description
- 1_Deterministic stock assessment
- 2_Incorporating uncertainty in stock assessment
- 3_Estimating data-poor life history
- 4_Generating ACLs
- CIE Reef Fish Life History Information_Humphreys_FINAL_08Sept2015
- CIE_Intro_Sep_2015
- Dunlap_CIE reef fish lengths mgmt overview
- Flow chart
- Hawaii Commercial Data_MKLowe_2015_0908_NadonCIE
- Williams CRED Reef Fish Visual Survey Program Overview

Further presentations and responses to requests:

1_Deterninistic stock assessment

2_Incorporating uncertainty in stock assessment

4_Generating ACLs

5_New graph

Division of Aquatic Resources _ Marine Fishes and Other Vertebrates

APPENDIX 2

Attachment A: Statement of Work for Dr. Kevin Stokes

External Independent Peer Review by the Center for Independent Experts

Review of length-based stock assessment methods for coral reef fish stocks in Hawaii and other U.S. Pacific territories

Scope of Work and CIE Process: The National Marine Fisheries Service's (NMFS) Office of Science and Technology coordinates and manages a contract providing external expertise through the Center for Independent Experts (CIE) to conduct independent peer reviews of NMFS scientific projects. The Statement of Work (SoW) described herein was established by the NMFS Project Contact and Contracting Officer's Technical Representative (COTR), and reviewed by CIE for compliance with their policy for providing independent expertise that can provide impartial and independent peer review without conflicts of interest. CIE reviewers are selected by the CIE Steering Committee and CIE Coordination Team to conduct the independent peer review of NMFS science in compliance the predetermined Terms of Reference (ToRs) of the peer review. Each CIE reviewer is contracted to deliver an independent peer review report to be approved by the CIE Steering Committee and the report is to be formatted with content requirements as specified in **Annex 1**. This SoW describes the work tasks and deliverables of the CIE reviewer for conducting an independent peer review of the following NMFS project. Further information on the CIE process can be obtained from www.ciereviews.org.

Project Description: The Pacific Islands Fisheries Science Center (PIFSC) is conducting stock assessments on exploited coral reef fish species in Hawaii, American Samoa, Guam, and the Commonwealth of the Northern Mariana Islands which are listed in the Western Pacific Regional Fishery Management Council's Fishery Ecosystem Plans. These stocks are generally classified as data-poor due to a lack of reliable, long-term, catch and fishing effort data. However, some parsimonious assessment models rely on more easily obtainable length composition data and certain key population demographic parameters related to growth, maturity, and longevity. PIFSC scientists have been implementing an approach that uses the average length in the exploited phase of the population (L_{bar}) to obtain an estimate of total and fishing mortality rates for coral reef fish stocks (Beverton & Holt 1956; Ehrhardt & Ault 1992). These rates, combined with population demographic parameters, are used in numerical population models to obtain stock sustainability metrics (e.g., spawning potential ratio, F/F_{msy}, B/B_{msy}; see Ault et al. 1998, 2008). Acceptable Biological Catches (ABCs) can be generated by obtaining recent total catch estimates and specifying new ABCs based on the results of the population sustainability analyses. Furthermore, a novel meta-analytical approach using stochastic simulations was developed at PIFSC to obtain demographic parameter estimates for species with even less data than data-poor species ("data-less" species). These scientific analyses have not previously been applied for management purposes in the Pacific Islands Region, so there is a need to conduct an independent peer review of the analyses to improve the scientific basis for management.

The Terms of Reference (ToRs) of the peer review are attached in **Annex 2**. The tentative agenda of the panel review meeting is attached in **Annex 3**.

Requirements for CIE Reviewers: Three CIE reviewers shall conduct an impartial and independent peer review in accordance with the SoW and ToRs herein. CIE reviewers shall have working knowledge and recent experience in the application of: general fisheries stock assessment, familiarity with length and age-based fishery models, and data-poor approaches to conducting stock assessments.

Each CIE reviewer's duties shall not exceed a maximum of 14 days to complete all work tasks of the peer review described herein.

Location of Peer Review: Each CIE reviewer shall conduct an independent peer review during the panel review meeting scheduled in Honolulu, HI during September 8^{th}-11^{th}, 2015.

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

Prior to the Peer Review: Upon completion of the CIE reviewer selection by the CIE Steering Committee, the CIE shall provide the CIE reviewer information (full name, title, affiliation, country, address, email) to the COTR, who forwards this information to the NMFS Project Contact no later the date specified in the Schedule of Milestones and Deliverables. The CIE is responsible for providing the SoW and ToRs to the CIE reviewers. The NMFS Project Contact is responsible for providing the CIE reviewers with the background documents, reports, foreign national security clearance, and other information concerning pertinent meeting arrangements. The NMFS Project Contact is also responsible for providing the Chair a copy of the SoW in advance of the panel review meeting. Any changes to the SoW or ToRs must be made through the COTR prior to the commencement of the peer review.

Foreign National Security Clearance: When CIE 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 CIE reviewers who are non-US citizens. For this reason, the CIE 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 before the peer review in accordance with the NOAA Deemed Export Technology Control Program NAO 207-12 regulations available at the Deemed Exports NAO website:
http://deemedexports.noaa.gov/
http://deemedexports.noaa.gov/compliance_access_control_procedures/noaa-foreign-national-registration-system.html

Pre-review Background Documents: Two weeks before the peer review, the NMFS Project Contact will send (by electronic mail or make available at an FTP site) to the CIE reviewers the necessary background information and reports for the peer review. In the case where the documents need to be mailed, the NMFS Project Contact will consult with the CIE Lead Coordinator on where to send documents. CIE reviewers are responsible only for the pre-review documents that are delivered to the reviewer in accordance to the SoW scheduled deadlines specified herein. The CIE reviewers shall read all documents in preparation for the peer review.

Material to be provided: Published scientific papers describing the approach and how it was applied in Florida (Ault et al. 1998) and Puerto Rico (Ault et al. 2008). Submitted paper applying this method to Hawaii (Nadon et al. 2015). Two un-published papers explaining and testing a new approach to obtain missing life history parameters (Nadon & Ault 2015a, 2015b). Other articles describing length-based methods or examining certain aspects of this approach (Ehrhardt & Ault 1992; Ault et al. 2005; Gedamke & Hoenig 2006).

Ault, J. S., J. A. Bohnsack, and G. A. Meester. 1998. A retrospective (1979-1996) multispecies assessment of coral reef fish stocks in the Florida Keys. Fishery Bulletin 96:395–414.
Ault, J. S., S. G. Smith, and J. A. Bohnsack. 2005. Evaluation of average length as an estimator of exploitation status for the Florida coral-reef fish community. ICES Journal of Marine Science 62:417–423.
Ault, J. S., S. G. Smith, J. Luo, M. E. Monaco, and R. S. Appeldoorn. 2008. Length-based assessment of sustainability benchmarks for coral reef fishes in Puerto Rico. Environmental Conservation 35:221–231.
Beverton, R. J. H., and S. J. Holt. 1956. A review of methods for estimating mortality rates in exploited fish populations, with special reference to sources of bias in catch sampling. Rapports et proces-verbaux des reunions du Conseil International pour l'Exploration de la Mer 140:67–83.
Ehrhardt, N. M., and J. S. Ault. 1992. Analysis of two length-based mortality models applied to bounded catch length frequencies. Transactions of the American Fisheries Society 121:115–122.
Gedamke, T., and J. M. Hoenig. 2006. Estimating mortality from mean length data in nonequilibrium situations, with application to the assessment of goosefish. Transactions of the American Fisheries Society 135:476–487.
Nadon, M. O., and J. S. Ault. 2015a. A stepwise stochastic simulation approach to obtain missing life history parameters for data-poor fisheries. Unpublished:1–34.
Nadon, M. O., and J. S. Ault. 2015b. Assessment of data-poor Hawaiian coral reef fish populations using life history parameters obtained through a stepwise stochastic simulation approach. Unpublished:1–31.
Nadon, M. O., J. S. Ault, I. D. Williams, S. G. Smith, and G. T. DiNardo. 2015. Length-based assessment of coral reef fish populations in the Main and Northwestern Hawaiian Islands. Unpublished:1–28.

Panel Review Meeting: Each CIE reviewer shall conduct the independent peer review in accordance with the SoW and ToRs, and shall not serve in any other role unless specified herein.

Modifications to the SoW and ToRs can not be made during the peer review, and any SoW or ToRs modifications prior to the peer review shall be approved by the COTR and CIE Lead Coordinator. 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 NMFS Project Contact is responsible for any facility arrangements (e.g., conference room for panel review meetings or teleconference arrangements). The NMFS Project Contact is responsible for ensuring that the Chair understands the contractual role of the CIE reviewers as specified herein. The CIE Lead Coordinator can contact the Project Contact to confirm any peer review arrangements, including the meeting facility arrangements.

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

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

1) Conduct necessary pre-review preparations, including the review of background material and reports provided by the NMFS Project Contact in advance of the peer review.
2) Participate during the panel review meeting in Honolulu, HI, from September 8th-September 11th, 2015, and conduct an independent peer review in accordance with the ToRs (**Annex 2**).
3) No later than September 25, 2015, each CIE reviewer shall submit an independent peer review report addressed to the "Center for Independent Experts," and sent to Dr. Manoj Shivlani, CIE Lead Coordinator, via email to mshivlanim@ntvifederal.com, and Dr. David, Die, CIE Regional Coordinator, via email to ddie@rsmas.miami. Each CIE report shall be written using the format and content requirements specified in Annex 1, and address each ToR in **Annex 2**.

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

August 4, 2015	CIE sends reviewer contact information to the COTR, who then sends this to the NMFS Project Contact
August 25,2015	NMFS Project Contact sends the CIE Reviewers the pre-review documents
September 8 – September 11, 2015	Each reviewer participates and conducts an independent peer review during the panel review meeting
September 25, 2015	CIE reviewers submit draft CIE independent peer review reports to the CIE Lead Coordinator and CIE Regional Coordinator
October 9, 2015	CIE submits CIE independent peer review reports to the COTR
October 16, 2015	The COTR distributes the final CIE reports to the NMFS Project Contact and regional Center Director

Modifications to the Statement of Work: This 'Time and Materials' task order may require an update or modification due to possible changes to the terms of reference or schedule of milestones resulting from the fishery management decision process of the NOAA Leadership, Fishery Management Council, and Council's SSC advisory committee. A request to modify this SoW must be approved by the Contracting Officer at least 15 working days prior to making any permanent changes. The Contracting Officer will notify the COTR within 10 working days after receipt of all required information of the decision on changes. The COTR can approve changes to the milestone dates, list of pre-review documents, and ToRs within the SoW as long as the role and ability of the CIE reviewers to complete the deliverable in accordance with the SoW is not adversely impacted. The SoW and ToRs shall not be changed once the peer review has begun.

Acceptance of Deliverables: Upon review and acceptance of the CIE independent peer review reports by the CIE Lead Coordinator, Regional Coordinator, and Steering Committee, these reports shall be sent to the COTR for final approval as contract deliverables based on compliance with the SoW and ToRs. As specified in the Schedule of Milestones and Deliverables, the CIE shall send via e-mail the contract deliverables (CIE independent peer review reports) to the COTR (William Michaels, via William.Michaels@noaa.gov).

Applicable Performance Standards: The contract is successfully completed when the COTR provides final approval of the contract deliverables. The acceptance of the contract deliverables

shall be based on three performance standards:
(1) The CIE report shall completed with the format and content in accordance with **Annex 1**,
(2) The CIE report shall address each ToR as specified in **Annex 2**,
(3) The CIE reports shall be delivered in a timely manner as specified in the schedule of milestones and deliverables.

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

Support Personnel:

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

Manoj Shivlani, CIE Lead Coordinator
Northern Taiga Ventures, Inc. Communications
10600 SW 131st Court, Miami, FL 33186
mshivlani@ntvifederal.com Phone: 305-968-7136

Key Personnel:

NMFS Project Contact:

Marc Nadon
Stock Assessment Scientist
Fisheries Research and Monitoring Division
National Marine Fisheries Service
Pacific Islands Fisheries Science Center
1845 Wasp Boulevard, Bldg. #176
Honolulu, Hawaii 96818
marc.nadon@noaa.gov
808-725-5317

Annie Yau
Annie.yau@noaa.gov
808-725-5350

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

1. The CIE independent report shall be prefaced with an Executive Summary providing a concise summary of the findings and recommendations following Annex 2 Terms of Reference questions.

2. The main body of the reviewer report shall consist of a Background, Description of the Individual Reviewer's Role in the Review Activities, Summary of Findings for each ToR in which the weaknesses and strengths are described, and Conclusions and Recommendations in accordance with the ToRs.

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

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

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

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

3. The reviewer report shall include the following appendices:

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

Annex 2: Terms of Reference for the Peer Review

Review of length-based stock assessment methods for coral reef fish stocks in Hawaii and other U.S. Pacific territories

1. Review the assessment methods used: determine if they are reliable, properly applied, and adequate and appropriate for the species, fisheries, and available data considering that the data itself have been accepted for management purposes.

2. Evaluate the implementation of the assessment methods: determine if data in its current form are properly used, if choice of input parameters seems reasonable, if models are appropriately specified and configured, assumptions are reasonably satisfied, and primary sources of uncertainty accounted for.

3. Comment on the scientific soundness of the estimated population benchmarks and management parameters (e.g., spawning potential ratio, F/F_{msy}, B/B_{msy}, stock status) and their potential efficacy in addressing the management goals stated in the relevant FEP or other documents provided to the review panel.

4. Determine whether the results (such as SPR-based reference points, stock status) in their current form from the assessment methods can be used for management purposes without further analyses or changes considering that the data itself have been accepted for management purposes.

5. Suggest research priorities to improve our understanding of essential population and fishery dynamics necessary to formulate best management practices. Comment on alternative data sources and modeling.

6. Draft a report of the WPSAR Panel conclusions and findings, addressing each Term of Reference.

APPENDIX 3
PERTINENT INFORMATION FROM THE REVIEW

1) Participation List

Tuesday, September 8th 2015

- Marc Nadon, NOAA Pacific Islands Fisheries Science Center (PIFSC)
- Cathy Dichmont, CIE reviewer
- Kevin Stokes, CIE reviewer
- Graham Pilling, CIE reviewer
- Ivor Williams, PIFSC
- Beth Lumsden, PIFSC
- Jon Brodziak, PIFSC
- Robert Humphreys, PIFSC
- Kimberley Lowe, PIFSC
- Adel Heenan, PIFSC
- Matt Dunlap, NOAA Pacific Islands Regional Office (PIRO)
- Adel Heenan, PIFSC

Wednesday, September 9th 2015

- Marc Nadon, PIFSC
- Cathy Dichmont, CIE reviewer
- Kevin Stokes, CIE reviewer
- Graham Pilling, CIE reviewer
- Ivor Williams, PIFSC
- Annie Yau, PIFSC
- Christopher Boggs, PIFSC
- Matt Dunlap, PIRO
- Marlowe Sabater, Western Pacific Fishery Council

Thursday, September 10th 2015

- Marc Nadon, PIFSC
- Cathy Dichmont, CIE reviewer
- Kevin Stokes, CIE reviewer
- Graham Pilling, CIE reviewer
- Christopher Boggs, PIFSC
- Annie Yau, PIFSC

Friday, September 11th 2015

- Marc Nadon, PIFSC
- Cathy Dichmont, CIE reviewer
- Kevin Stokes, CIE reviewer
- Graham Pilling, CIE reviewer
- Annie Yau, PIFSC
- Christopher Boggs, PIFSC

2) Final Agenda

Agenda

Review of length-based stock assessment methods for coral reef fish stocks in Hawaii and other U.S. Pacific territories

University of Hawaii at Manoa, Hemenway Hall, Room 208

8-11 September 2015

Tuesday September 8 (9:00 am – 4:00 pm)

1. Introduction (Brodziak)
2. Objectives and Terms of Reference (Brodziak)
3. Fishery (Nadon)
4. Data
    - Commercial data (Lowe)
    - Diver survey data (Williams)
    - Age & Growth (Humphrey)
5. Management process (Makaiau)
6. Review of Stock Assessment (Nadon)

Wednesday September 9 (9:00 am – 4:00 pm)

6. Review of Stock Assessment - continued (Nadon)

Thursday September 10 (9:00 am – 4:00 pm)

7. Continue Assessment Review (1/2 day)
8. Panel discussions (Closed)

Friday September 11 (9:00 am – 4:00 pm)

9. Panel Discussions (1/2 day)
10. Present Results (afternoon)
11. Adjourn