PIFSC Stock Assessment Review

Reviewer 2

Summary
This review examines the Pacific Island Fisheries Science Center's fisheries stock assessment program. There are some significant challenges in conducting stock assessments in this regions due to the large geographic area associated with the Main Hawaiian Islands the the numerous territories. There are three general stock assessment divisions: (1) highly migratory species that are of high value, fisheries are prosecuted in both domestic and international waters, and employs relatively few people, (2) insular bottomfish species that are the target of a few hundred commercial fishermen and represent a small fraction of the total value, and (3) coral reef species that are much smaller in scale, but of great cultural importance to smaller recreational fisheries for local island subsistence. There are many strengths of the existing stock assessment programs, and many weakness that could be addressed with additional data, additional quantitative technical staff, and a more streamlined process for conducting assessments on a regular schedule. Perhaps the biggest challenge facing the stock assessment program at PIFSC is attracting and retaining qualified technical staff who are capable of conducting stock assessments, identifying significant issues that are of management concern, and conducting the appropriate research to address quantitative needs for PIFSC and decision makers. Mentoring and training of the young fisheries scientist should be given high priority within PIFSC. Emphasis should be placed on moving away from simple production models towards structured models that allow for incorporation of other existing data that is not being used to its fullest capacity. I also comment on staffing issues, training, and mentorship within the stock assessment division.

GENERAL OBSERVATIONS

1. Scientific & technical approach to fisheries stock assessment modeling.

In comparison to many of the other National Science Centers, there are many novel aspects of the data streams, assessment methods and technical approaches being implemented in the Pacific Island Fisheries Science Center. The technical approaches that are currently fall into three general categories of data rich, moderate data, and data poor species. This categorization also happens to align with highly migratory species, bottomfish complexes and coral reef assemblages respectively. A variety of methods are employed ranging from relatively simple analyses of size composition data and inferring total mortality rates and spawning potential ratios, to highly sophisticated statistical models that integrate multiple sources of information.

One of the major goals in the scientific approach to stock assessment modeling is to elicit the underlying production function of a unit stock. All of these models have two key parameters that define stock productivity in relation to its relative abundance, and the global scale of the population. These are also defined as the maximum intrinsic rate of growth and the carrying capacity in a simple production model, or the steepness of a stock-recruitment relationship and the unfished biomass in an age-structured model. These two key leading parameter are generally negatively correlated with each other (e.g., the stock is either large and unproductive, or small and very productive). This negative correlation is particularly problematic in cases

where the available data only consists of a relative abundance index that is monotonically decreasing (e.g., deep-7 bottomfish), increasing, or has no trend. This confound exists regardless if the relative abundance index is fisheries-dependent or fisheries-independent. Even in cases where the assessments appear to be data rich, these one-way trip data sets can be information poor with respect to two key population parameters that define the underlying production function.

In cases where only catch and CPUE data exits, and the CPUE data represent a one-way trip, it is very unlikely that the confounding between stock productivity and global scaling can be resolved from these data alone. In order to proceed with conducting a stock assessment, assumptions in the form of fixing model parameters, or using an informative prior distribution, are necessary to proceed with a Bayesian or penalized likelihood assessment. The development of prior distributions should be independent of the data being used in the assessment, and both statistics (e.g., mean and variance) of the prior density and posterior density should be expressed, or shown visually as density plots with the prior and posterior on the same axis. In a full Bayesian assessment, it is also desirable to examine the outcome where P(øIx) = 1; or effectively the likelihood of the data is turned off. This provides reviewers opportunities to understand how informative proper prior distributions influence the model outcome and management advice.

The use of relatively simple production models is a suitable analytical method and has advantages over more explicit age or size-structured approaches in terms of information requirements. However, there are also some drawbacks to these approaches, the biggest is the limitations on being able to incorporate other relatively 'cheap' auxiliary information such as observed changes in mean size of the catch. Also, simple production models are known to have significant lags and the length of the lag is proportional to the age-at-recruitment to the fishery; they also perform poorly in short-term predictions. For example, when the stock is increasing the production models tend to under-estimate biomass and when the stock is decreasing production models tend to over-estimate biomass. This is particularly worrisome when the stock sizes are decreasing and it will take of number of years to discover overfishing is occurring as the stock fails to respond due to an upward bias in the catch advice.

It is possible to use and parameterize age-structured models to accomplish the same task as a simple production model in the absence of detailed information (e.g., assume knife-edge selectivity and maturity which is equivalent to the assumptions of a production model). This allows for direct integration of life-history information, rather than developing an informative prior to approximate life-history information. Given that estimates of natural mortality rates are already being used to develop priors for the maximum intrinsic rate of growth, the only other additional information required is the relative weight-at-age (or growth parameters) to translate numbers-at-age into biomass-at-age.

I commend the PIFSCs efforts towards addressing the data poor coral reef assemblages. The simulation testing of the methods being developed by Marc Nadon is a compelling demonstration that these methods can work provided the assumptions are met. Moreover, the simulation testing allows for exploration of outcomes if some of the more tenuous assumptions are not met, and exactly what additional information is required to shore-up those assumptions. The estimates of total mortality rate (Z) and Spawning Potential Ratio (SPR) based on size composition data is an excellent first step in determining the relative depletion of the stock and stock status relative to some arbitrary proxy for what would be considered optimal depletion

(e.g. $SPR_{30\%}$). The next major hurdle will be translating this information in such a way that it can be used to set Annual Catch Limits (ACLs) for each group or assemblages of coral reef species.

2. Process for conducting stock assessments.

For the insular stock assessments, there is no formal terms of reference that clearly states what information should be provided in order to make informed management decisions that explicitly states how to deal with uncertainty. Developing a detailed terms of reference (TOR) with the WPFMC, PIFSC, and the SSC should receive a high priority, and should also be an iterative process involving all parties. Many of the other councils have already articulated the essential elements that should be included in the TOR and these could easily be adopted immediately, or quickly modified to address local PIFSC needs.

Stock assessment methods have evolved considerably in recent decades with advances in computers and software for estimating parameters in non-linear models. I suspect these trends will continue. Moreover, new data and structural changes to the model also evolve over time, and the stock assessment documents should represent a written record of this evolution. For a first time reviewer, or new analyst, this history is an important lesson on how the process has evolved to its current state. That being said, all of the stock assessment documents should include sections that reflect the comments and the analysts responses to the various Plan Teams, SSC, and reviews through the Western Pacific Stock Assessment Review (WPSAR) and CIE reviews.

The assessment process and TOR should also include a detailed description of the available data used in the assessment along with other data that are available that might inform the assessment both quantitatively or qualitatively (anecdotal information from commercial and recreational fishermen).

3. Peer review process.

Peer review stock assessments and stock assessment methods is necessary to ensure the assessments are up to current status with respect to acceptable statistical methods and to ensure that it is responsive the the councils needs and are suitable for setting ACLs. The new WPSAR process is in its infancy and should mature to a more efficient stage as the process develops with more and more case studies. Currently the process relies on review from the Center of Independent Experts (CIE) and the Council's SSC, and is led by a steering committee with members from the PIFSC, PIRO and Council staff.

The specific terms of reference and each of the stock assessment documents produced by the PIFSC should include sections outlining comments from the plan teams and comments from reviews conducted by the Center for Independent Experts (CIE). The history of these comments should be maintained in each of the benchmark assessments as well as the updates based on the benchmark model. Currently there is no funding for the peer review process. It would be preferable for CIE reviewers to be directly involved face to face in the review process, rather than a written desk review from afar. In addition, the peer review process should make

the source code and data available to the reviewer such that the results of the assessment can be repeated independently by the external reviewer. Often a written description of the model code is insufficient to repeat the results and moreover, often the code does not exactly reflect what is in the written description. This is an essential error checking process.

There is an established history of CIE reviewers' coming in and making major changes to the data streams, structural changes to the model, or other modifications that can have profound impacts on the process. Normally, the analyst would have much more experience with the history and legacy of the data. Formal protocols must be established in cases where there is major disagreements between the analysts and peer reviewers if such circumstances were to arise.

4. Organization and priorities

There are a large number of of species complexes that are in the WPFMC fisheries management plan that require ACLs, and there are currently ALCs in place for all of these groups, some of which are simply *ad hoc* until a new process is developed to evaluate the current stock status. A priority list for each of the highly migratory stocks, bottomfish, and reef fish complexes need to be established, along with a schedule of how often these assessments and ACLs are updated for each of these groups. The priority list should also take into consideration the extent to which these fisheries are occurring in Federal waters and should be addressed directly by PIFSC, and which should be addressed by RFOs and state agencies.

5. Accomplishments.

There has been some recent turnover in staffing stock assessment positions at PIFSC and the center has been successful at recruiting a number of young, early career scientists, who are very talented and have demonstrated great potential with creative ideas for dealing with data poor species. There are also established connections with the University of Hawaii and other off island institutions that should help will filling in missing life-history information via Masters student projects looking at growth and maturity information for many of the insular species. These collaborations will require some financial support and mentoring from the PIFSC.

The local Universities, however, do not have quantitive fisheries scientist who are familiar with the quantitative tools necessary to conduct stock assessments. The newly hired quantitative ecologist at University of Hawaii (Prof. Eric Franklin) would be capable of jointly supporting students in fisheries science but will require direct mentoring from experience PIFSC staff. However, careful consideration must be given to Prof. Franklin with respect to his workload and ability to supervise students. At best I could imagine, in addition to his own research interest, that he could accommodate one new Masters student per year, and acquisition of life-history data for say 10-species could table unto a decade or more. This may not be the most efficient use of resources.

6. Communication and dissemination of scientific advice.

Communication and dissemination of scientific advice is mostly in the form of reports to the council and reports to RFO's. There are a large number of species in the bottomfish complexes, and even more in the coral reef fisheries. The current assessment staff of roughly 7 people is

producing on average about 2-4 assessments for all HMS, Bottomfish and coral reef species combined. Improved communication between field staff, people who assemble and process the necessary data should be improved to streamline the process of turning out assessments. The development of a formal terms of reference and an assessment schedule with council and PIRO would also improve communication streams between the various agencies and increase assessment activities and output.

7. Opportunities for improvement.

There are many opportunities to improve the current assessments that are now being conducted at PIFCS via incorporating addition composition information (size or age composition). These additional data can provide additional information on recruitment variability or relative cohort strengths, size- or age-based selectivity, estimates of total mortality rates conditional on assumptions about selectivity. The current use of production models assume that all fish recruit at age-1 and all fish are sexually mature at age-1. The use of an age-structure model would allow for additional sensitivity analyses on the effects of lags in the model. Independent estimates of total mortality would vastly improve estimates of population scale given that many of the relative abundance indices in the bottomfish fisheries consist of one-way trips. The integration of age- or size-composition data would improve estimates of population scale via catch curve type information.

For the coral reef fisheries and assessments, integration of creel survey data for estimates of total catch and absolute abundance estimates based on data from the Reef Fish Visual Survey Program (RAMP) should be integrated into the models being developed by Marc Nadon to take the approach to the next step and provide information necessary to set ABCs and ACLs.

Due to the many novel fisheries and challenges associated with data poor fisheries, the immediate data available is not amenable to may of the tools available in the NOAA toolbox. There are significant opportunities to collaborate with academics and other Centers who face the same data poor challenges. Developing and simulation testing of new analytical tools could place the PIFSC at the forefront of developing new methods, but would require additional analytical support in the form of scientific programmers who are skilled in a diverse set of computer languages to support fisheries scientists in analyzing data and disseminating information to the council, managers, and the general public.

PIFSC is essentially groping for a new technology to come up with a fisheries independent survey (FIS) for the bottomfish stocks and implement stock assessment methods and data streams that are consistent with methods used by other centers that have access to more relative abundance and composition information. Since the original discussions for an FIS that began in 2004, nearly 10 years has passed and there is still no single technology that has been identified to come up with a reliable index. At present there are at least four different technologies being tested, all competing for the same limited financial resources. Moreover, there are no specific stock assessment objectives for the survey program in place. For example, one specific objective would be to design a program that is able to detect a 20% change in biomass with a probability of 0.95. Given such a clearly stated objective, and the preliminary data that has been collected thus far, it should be possible to determine what sort of resources are required to achieve such an objective.

The PIFSC is doing some interesting research with respect to model selection based on statistical criterion on how well alternative models fit simulated data using self-testing and cross-model testing. This is an important step and process for understanding how alternative models, model averaging, and statistical selection criterion (e.g., DIC) perform in correctly identifying a suitable suite of models. However, model selection and multi-model inference should be based both on statistical performance as well as policy performance. For example, a model with biased parameter estimates may perform just as well in the policy arena as a model with unbiased parameter estimates. How well the models perform from a policy perspective should be given more emphasis in stock assessment research programs. Management Strategy Evaluation is a good tool for moving research in this direction, where the first priority is to set fisheries management objectives and performance measures; only from there is it possible to explore how well alternative procedures will perform with respect to policy.

KEY FINDINGS AND RECOMMENDATIONS

The following bullets, in no particular order of importance, highlight some of my key findings and recommendations for conducting stock assessments, reviewing assessments, and ensuring that development of new technologies and models are done in a consistent and timely fashion and contribute to the overall assessment goals.

- For the hawaiian bottomfish complex there is information on mean weight of the catch and size composition that are not being used in the stock assessments. Changes in the mean weight of the catch would provide additional information about the relative contributions of fishing mortality versus recruitment and if integrated with trends in the CPUE data would provide more reliable estimates of the global scaling parameters that are currently informed via necessary subjective decisions made by the analysts. Incorporating size-composition would improve estimates of global scaling and stock size and provide insights on how the fishery is being prosecuted (e.g., observing increases in mean size with declining CPUE would be indicative of fisheries expansion, or changes in selectivity towards larger more valuable fish that are less abundant).
- Adopt the use of age-structured models (either explicit models or delay-difference models) to reduce the potential for lags, and allow for incorporation of other auxiliary data such as mean weight of the catch, or size-composition information. Even in the absence of composition information, these models are still amenable for use with catch and CPUE data only, are more parsimonious, and permit important sensitivity analyses.
- Winbugs is very limited tool for developing structurally complex models. More modern statistical software should be adopted for use in stock assessment (e.g., AD Model Builder, or Template Model Builder. … even Excel!).
- All of the assessment tools should undergo extensive simulation testing to demonstrate that the available data are informative about the estimated parameters, and if not what sort of assumptions or prior information is required to address these shortcomings.
- For assessments in which the model is being fitted to data, statistics informing that the model is full rank should be reported in the assessment documents. This can be determined by singular value decomposition of the hessian (or covariance) matrix (or the *.eva) files from standard ADMB output.

- For assessments that use prior distributions, statistical descriptions (i.e., mean and variance) of the prior density and the posterior density should be given in tables, as well as, density plots where both the prior and posterior distributions are shown on the same panel.
- In full Bayesian assessments, scenarios where the log likelihood of the data is set equal to 0 should be presented to allow reviewers and decision makers to discern how informative proper priors influence management advice.
- Conduct simulation testing of the anticipated data to be obtained from BotCam survey results to justify the expense of this fisheries independent survey. If these data cannot be used to generate estimates of absolute abundance, then specific questions should focus on how much contrast in the trends are required to resolve correlations in key model parameters (e.g, r and k, or steepness and Bo), and what sort of coefficient of variation is required to detect these trends with a probability of say 0.95? Should experimental fishing be invoked to generate the necessary contrast?
- Research on stock assessment models has only focused on statistical performance of simple production models (i.e., how well the model can recover the true parameter values). A more important issue is how well these models perform in the policy arena. More research and exploration of policy performance from alternative models should be the focus, rather than judging solely on statistical performance. Many of the statistical properties of simple production models are well known and are published in the primary literature and should probably be used conservatively.
- Periodically make concerted efforts to obtain composition data (preferably age-composition data) to better inform estimates of selectivity, growth, and total mortality rates. For long-lived fish, the sampling frequency would be less frequent than short-lived fish. Size-composition data is relatively cheap to collect, and combined with periodic age-composition samples would vastly improve estimates of global scaling in the population models. Yes there are additional assumption regarding, but these same assumptions are implicit in simple production models.
- For the coral reef fisheries, the new methods being developed based on size-composition sampling only are only informative about the relative stock status. Global scaling information such as catch data, or independent estimates of absolute biomass are required to translate this information into ABCs and ACLs.
- Develop a formal terms of reference for stock assessment documents (which should be the same for benchmark and update assessments) that clearly states what information should be contained in the executive summary, documentation of the stock assessment models, parameters, parameter bounds, and standard errors etc., response to Plan Team Comments, response to SSC comments, and response to CIE review comments.
- Establish protocols for contesting reviews.
- Peer review panel should serve 3-5 year terms, with turnover of one new individual each year to keep a fresh set of eyes on the stock assessment process and input data streams.
- The PIFSC should hire a scientific programmer to help deal with processing of data and developing efficient tools for use in stock assessment, decision making, and policy exploration using Management Strategy Evaluation and helping staff develop more modern tools to address the complex problems in the Hawaiian Islands and territories.
- Find long-term solutions for reducing staff turnover in the Center. Living on Oahu is challenging, especially for young people who are about to start a family in a city with an

extremely hight cost of living. Recruiting and training residents of Hawaii and other territories may be one solution to this problem.

CONCLUSIONS

In conclusion, I would recommend that five major areas would lead to significant increases in productivity, quality, and inspiration in the PIFSC: (1) improved mentoring and training of young assessment scientist and data analysts, (2) integrate more modern assessment methods and continue to foster the creative ideas for data poor species, (3) define a clear terms of reference for cooperative collaborations between assessment scientist, data analysts, and council SSC and staff, (4) develop more scientific programming capacity within the Center, and (5) find a permanent solution for dealing with the recent high turnover rate of fisheries scientist.

Recently, there has been a large turnover is assessment staff and a retirement of senior staff have already started and more will happen in the near future. There are a number of young bright fisheries scientist that are now in house, and need proper mentoring with modern assessment tools and the notion that statistics performance is only one part of the equation for model selection. The critical components to model selection is how well it performs in the policy arena. The Center should try and recruit a Fisheries Scientist who is more experience in the area of Management Strategy Evaluation and the use of integrated models.

Of the two examples shown to the the review panel, only the Blue Marlin assessment model is capable of dealing with the complex nature of the multiple fisheries and data available for this stock. The Hawaiian bottomfish assessment model is very simple; albeit does contain a lot of data, the data that are used are information poor in terms of estimating key population parameters. Accommodation of additional data is possible, but the development of more structurally complex models is necessary.

Its challenging to operate in the dark. Working with council SSC, PIRO and PIFSC on developing a clear terms of reference is necessary, especially as a road map for junior scientists who are conducting their first assessments here.

Implementing new and novel stock assessment research here in the main Hawaiian Islands and territories is going to require more scientific programming support. Efficient tools for conducting multi-species assessment and management strategy evaluation are not yet widely available, and products such as stock synthesis are limited in capacity to conduct Bayesian assessments and efficient MSE work. I would highly recommend recruiting a scientific programmer with some fisheries related experience to aid in the development of new tools necessary for these challenges.

It is my opinion that the PIFSC faces many difficult challenges in the short-term the biggest of which is retaining the current young staff, creating a better work environment for these people, and fostering creativity to tackle difficult problems. The enthusiasm is there, but the mentoring is not. Relative to other Centers around the country the PIFSC is lags far behind in the available data, assessment models, and quantitative tools required to address the status of many of the insular stocks. More training on the modern tools that are available and moving beyond simple production models is necessary to bring the center up to the standards that are now in practice world wide.

These young scientist are the future of PIFSC, and passive aggressive and condescending comments from their mentors are not effective at fostering creative thinking and developing a fun place to work. I was witness to these behaviours during this review. This behaviour in the workplace should be dealt with strictly by the senior administration. I have no professional credentials for judging, creating, or fostering a productive work environment, but the three areas I would focus on for retaining staff: (1) strong leadership, (2) staff moral, and (3) fair and equitable compensation. Aggression in the workplace should not be tolerated.

In short, I foresee a real crisis in maintaining and staffing stock assessment and fisheries scientists at the PIFSC in the short-term. Globally, there is a shortage of people with the technical skills for conducting stock assessments and doing research in quantitative fisheries science; these young bright recruits could easily move to a better work environment so early in their careers. Having a continuous turnover in staff will lead to continued delays in progressing and developing the stock assessment programs and fisheries science at PIFSC.