Reviewer report for 2024 Japanese domestic stock assessments

December 2024

Peter T Kuriyama

NOAA Southwest Fisheries Science Center
8901 La Jolla Shores Drive
La Jolla, CA 92037
peter.kuriyama@noaa.gov

1. General Comments

First, I would like to thank the assessment presenters and authors. I think the review meeting went smoothly and the presentations responded well to my reviewer comments. I enjoyed my time learning about the Japanese stock assessment process and unique aspects of the data and management. I would also like to thank the meeting coordinators for organizing our travel and accommodation.

I think there may be benefit to having more time dedicated to review each stock assessment at the in-person meeting. I think the current format where the reviewers read the document and submit preliminary questions in preparation for the in-person review is adequate, but I do think there are a number of components to the assessments that the reviewers and scientists are unable to investigate further at the meeting. For example, for the NOAA stock assessments we will typically have a two-to-three-day meeting to review one stock assessment. This allows for reviewers to ask for model runs on day one that the assessment authors will evaluate overnight and present the following day. This iterative approach results in a fuller exploration of the information quality of the data and also the influence of assumptions on the model results. Within this review, the assessment authors and reviewers may agree to drop specific data sets or pool data sets (e.g., combine season-specific age-length keys into an annual age-length key) in the assessment. This poses additional logistical complexities and would involve a higher amount of investment, which may not be feasible.

I think a number of the assessments would benefit from having a more flexible modeling framework. Virtual population analysis is perhaps most appropriate for a stock like Pacific cod, Honshu Northern Pacific in which the number of age and length observations is high and there is a fishery-independent index of abundance and perhaps less appropriate for Japanese amberjack where there is high uncertainty in the age-length relationships. A statistical integrated catch-at-age modeling framework like Stock Synthesis would allow for a more robust exploration of the influence of these data uncertainties and assumptions on the results. Stock Synthesis can estimate parameters with a penalized likelihood in which the assessment authors could define a reasonable range of uncertainties around parameters (similar to what was done for Pointhead flounder), and this type of approach will likely be beneficial for a parameter like natural mortality rather than assuming it is fixed at one value. Stock Synthesis involves a fair amount of training time and expertise to develop a model, but once the model files are configured, it is relatively easy to modify model configurations like selectivity curves or stock-recruit relationships. I know that this type of shift will take time to implement, but I think an alternative modeling approach will facilitate model development and be a bit easier to document and present to reviewers.

To the extent possible, it will likely be beneficial to develop a more coordinated program for ageing otoliths. Some stocks like Pacific cod have an extremely high number of ages read each year, whereas other stocks have very few. I might argue that the number of Pacific cod otoliths of about 1,000 reads each year may be excessive and reducing the number to a couple hundred may be sufficient to characterize the year-to-year variability.

The specific numbers of age reads should be explored with additional studies. I think it would be beneficial to develop a more unified ageing program to the extent possible, especially given the importance of high quality and representative age data for virtual population analyses.

2. Japanese amberjack

The stock structure for Japanese amberjack seems to be less of an issue than I initially thought. The assessment authors discussed that catches of Amberjack are relatively small in Taiwan in China. The presentation did mention that there are some tagging data available, and if these data are available, the assessment authors should explore model sensitivities to alternative population distribution scenarios. For example, say that the tagging data indicate that 10% of Japanese amberjack move outside of the distribution boundaries defined in the assessment. A model sensitivity might assume that recorded catches represent 90% of the "true" catch amounts and increase the catch values in the assessment. This model sensitivity would likely just highlight the degree of impact of alternative distribution assumptions on the stock assessment results.

The catch data used in the Japanese amberjack assessment comprise three species: Japanese amberjack (buri), yellowtail amberjack (hiramasa), greater amberjack (kanpachi). Japanese amberjack was reported to account for more than 90% of the amberjack landings, and the figures on (Slide 6 of the presentation) indicated that there are year-specific estimates of catch by species. One recommendation for the future is to get better estimates of the Japanese amberjack only catch for use in the next assessment.

Another recommendation is to evaluate a model sensitivity run in which the analysts calculate the average percentages of Japanese Amberjack and non-Japanese Amberjack catches based on available data. Then these average percentages could be applied to the catch time series used in the stock assessment. This model sensitivity will quantify the impact of improving accounting for the catch numbers.

The mojako are an interesting component of the data/model for Japanese amberjack. I agree with the authors decision to include these data in the stock assessment. It is important to include all sources of fishing mortality in the stock assessment. One model sensitivity in the future would be to evaluate the impact of assuming natural mortality (M) of 0.6 for mojako and 0.3 for 7–12-month-old Japanese amberjack. I might think that the early and late mojako M values might be higher for these age-0 fish.

Slide 10 of the assessment presentation described the recreational catch data available and the model sensitivities to including these data in the assessment. I appreciate the model sensitivity and recommend that the authors include additional justification for not including these data in the stock assessment.

The generation of catch-at-age data is a key component of virtual population analysis. I am still a bit unclear on the process of generating the catch-at-age data. My understanding of the process is that fish are landed at ports and sorted into commercial size categories (for example 2-4 kg are small buri, 4-7 kg are medium buri, etc). The total catch associated with each size class are then reported by port. The Watari *et al*. 2019 paper has estimates age-length and age-weight relationships from different regions of Japan for the years 2002-2017 (shown in Slide 25 of the presentation). These estimates can be broken out into weight-age classes by month, and these are the values that were used to convert the commercial size category data into ages (Slide 28 of the presentation).

A longer-term goal is to collect more specific data regarding the age-length and age-weight relationships. This would require a large change in the data collection methods and likely will not change by the next benchmark assessment. This could include a country-wide sampling program where scientists in each prefecture could be trained to sample the landed fish and collect length and weight measurements. Otoliths or scales could be collected and used to get age estimates for the fish. If these methods are someday adopted, a VPA or cohort analysis would be an appropriate modeling approach because there will likely be less uncertainty in these data.

In the meantime, an integrated statistical catch-at-age model (like Stock Synthesis) might be more appropriate for the available data. Specifically, this approach would allow for the uncertainty in the age-weight data to be accounted for in the model and incorporated in the stock assessment results. The commercial size category data would be applicable for input to Stock Synthesis. That is, all the data could be input to an integrated model and different age-length and length-weight relationships could be estimated. These might be time-varying with a different relationship for each year or perhaps each cohort. This approach would make the uncertainties in the data and their influence on the model results more explicit than they are currently presented.

3. Pacific cod, Honshu Northern Sea of Japan

Overall, I think a number of the assumptions in this assessment were justified given the available data. The stock seems to have a reasonable historical trend given that catch history, and the current stock status indicates that F is below F_{MSY} and B is above B_{MSY}. I think it may be beneficial to explore a model in the future that is more flexible than cohort analysis/virtual population analysis. Doing so might allow for the model to be extended with the inclusion of the index of abundance that does not have associated length composition information. I think that the data quality seems to be relatively high, and hopefully data collection and research on genetics, ageing, and potentially migrations will continue in the future.

The stock boundary seemed to be relatively precise, and the assessment presentation described the genetic monitoring and analysis that are used to identify the stocks. The stock boundaries seem reasonable given the available data, and there does not seem to be a large portion of the stock caught in neighboring regions (like Kyoto and Hyogo prefectures). Although it will likely be valuable to continue to evaluate genetic structure and potential mixing between different stocks because the current genetic structure does not correspond to management units.

Monitoring the genetic identities of spawning fish should be an area of future research (as discussed in the presentation). This monitoring may be important if the nearby stocks display some sort of density-dependent shift in habitat range. These fish may expand their distributions in periods of high abundance and contract under low abundance. The presentation discussed the Hokkaido and Mutsu-Bay stocks have been increasing since 2010 and are now at a record high level. This was highlighted in the presentation as an area of research that will continue into the future.

The number of individual fish aged per year (~500) was quite good, although the ageing only began in 2016. Efforts to measure the age of fish should be prioritized, particularly if cohort analysis will continue to be the primary assessment method. I believe the average age-length key was used to calculate numbers at age from the length compositions prior to 2016. It would be beneficial in the future to evaluate the model sensitivities to different age-length key assumptions. This might include using just the 2016 age-length key for the pre-2016 numbers-at-age calculations or to perhaps use the year-specific age-length key that most closely matches the year-specific length composition data. I think the current assumption of an average age-length key is perhaps the most reasonable assumption, but it would be helpful to better understand how sensitive the model results are to alternative assumptions.

I recommend exploring alternative modeling approaches for future versions of the assessment model. The use of an integrated statistical catch-at-age model (like Stock Synthesis) would allow for use of all the available data sets. The model could be expanded to begin in about 1970 (the beginning of index data) even without associated length compositions. It would also provide a means of evaluating the information quality of including or excluding the other available indices of abundance. Overall, this modeling approach will likely provide the ability to have more or less flexibility in the model. Specifically, the assumptions (for example different age-length keys, alternative selectivity forms, different stock-recruit relationships) and the impact of alternative assumptions would hopefully be more transparent.

4. <u>Pacific Cod, Honshu Northern Pacific</u>

I think this stock assessment was appropriately data rich for a VPA and the model estimates demonstrated reasonable stock dynamics. The presentation described that about 1,000 otoliths and about 40,000 length measurements were collected each year from a combination of fishery-independent surveys and market landings. The ageing uncertainty was described to be relatively low. Additionally, the number of samples was sufficient to develop half-year age-length keys and there was a fishery-independent index of abundance used to tune the VPA. Taken together, the assessment seemed to have reasonable results given the trajectory of the biomass and fishing mortality in the Kobe plots shown. Fishing mortality in 2019-2021 decreased below the F_{MSY} level, and the population had increases in spawning biomass and is approaching SB_{MSY}.

I think there are a number of interesting areas of future research for this stock. Many of these were described in the stock assessment presentation and were focused on the effects of the earthquake on different aspects of the population. One change was an increase in abundance from 2012-2016 that coincided with a decrease in growth and maturation rates. The data quality seems to be good enough that the changes in these processes are accounted for in the VPA, but shifting to an integrated stock assessment framework such as Stock Synthesis may offer some benefits.

The primary benefit of an alternative assessment approach would be to estimate parameters based on all the available data and allow exploration of alternative assumptions. In this data-rich case, the stock assessment authors could account for time-varying growth and time-varying fishery and survey selectivity (if supported by the data). Additionally, the authors could explore if stock-recruitment regimes or changes in survey catchability are best supported by the data. This isn't to say that a platform like Stock Synthesis will result in "better" or "truer" estimates, but rather this approach will facilitate model selection in which the authors can provide quantitative evidence for using one stock-recruit relationship over another for example. In other words, the model development process could be better documented in which the authors might provide support for the decisions around natural mortality, growth (age-length and length-weight relationships), recruitment, and maturity.

5. Pointhead Flounder, Northern Hokkaido

This assessment used a state-space surplus production model that estimated parameters with penalized likelihood. I think this approach was beneficial as it estimated uncertainties around the parameters (with influence from prior distributions) and incorporated these uncertainties in the model projections. This approach is good as the model results are less influenced by specific assumptions, which are often required in a VPA for example.

The available data from Hokkaido Research Organization seem to have interesting patterns. Unfortunately, the presentation discussed how these data are not available and this is unlikely to change in the future. Obviously, it would be useful to directly analyze the catch-at-age data in an assessment but there are a number of external difficulties that make the data unavailable.

That said, there seem to be some trends in the male and female catch data shared with the assessment authors. There may be some density-dependent growth processes occurring as evidenced by the 2000-2015 data in which there were low catches of males due to catch size restrictions. In this time period there was also an increase in abundance index 1, the VPA from Hokkaido Research Organization. Understanding these processes, to the extent feasible, would likely improve the assessment and forecast skill.

It may be justified to exclude the VPA index in the surplus production assessment model. The presentation described the justification for including the output from the Hokkaido Research Organization, but the model results seemed to not be particularly sensitive to exclusion of this index (the stock status is estimated to be the same in the “Surviving biomass only”, “CPUE only,” and “Base-case” models shown in slide 86 of the presentation).

I appreciate the authors’ responses to the reviewer questions about the specification of prior distributions around the parameters with standard deviations of 0.5 and 1 for the two candidate models. The presentation included a model without any prior distributions that had difficulty converging and finding robust estimates. The presentation also included a profile over the SDs for prior distributions of the parameters $q1$ and $\sigma1$. I think in this case, the use of the priors was justified as they resulted in model stability. However, I recommend the authors continue to refine the process of specifying these distributions as they had some influence on the assessment results.