Report of an Independent Peer Review of Four Sardine and Jack Mackerel Stock Assessments Conducted by the Japan Fishery Research and Education Agency (FRA) in 2021

By

Steven L. H. Teo
Research Fishery Biologist

Southwest Fisheries Science Center
National Oceanic and Atmospheric Administration (NOAA)
8901 La Jolla Shores Drive
La Jolla, California 92037-1508
USA

October 15th, 2021

Table of Contents

1. Summary

Assessments of two Japanese sardine (*Sardinops sagax melanostictus*) and two Japanese jack mackerel (*Trachurus japonicus*) stocks were conducted by the Japan Fishery Research and Education Agency (FRA) in 2021. These four stock assessments were on the: 1) Japanese sardine - Pacific (JS-P) ; 2) Japanese sardine - Tsushima (JS-T); 3) Japanese jack mackerel - Pacific (JM-P); and 4) Japanese jack mackerel - Tsushima (JM-T) stocks. An independent peer review of these assessments was organized by the Secretariat of Peer Review of the FRA in September 2021. This report is an independent review of the four assessments based on my personal opinions of the submitted documents and clarifications during the web meeting. The primary aims for this review were to evaluate if the current stock assessments were based on the best scientific information available (BSIA) and recommend improvements for future stock assessments.

All four assessments used virtual population analysis (VPA) models to estimate the population dynamics of the stocks during the historical period. Critically, a VPA model assumes that the catch-at-age is known without error. Therefore, this review focused on examining the data preparation and determining if this assumption was severely violated. Developing the catch-at-age time series depends largely on two components: 1) the total catch; and 2) the age composition of the catch. Based on the assessment documents and the web meeting, this review concluded that all four stock assessments severely violated the assumption that the catch-at-age was known without error, which likely led to large errors and biased results of the VPA models and assessments. Therefore, all four stock assessments did not meet the criteria of being BSIA.

Most importantly, the total catch used for three stock assessments (JS-P, JS-T, and JM-T) were considered inappropriate because substantial amounts of catch of the stocks by non-Japanese fisheries were excluded. The reasons for excluding these catches were that catch data from non-Japanese fisheries were often more uncertain and untimely compared to Japanese fisheries, and do not have corresponding age compositions. In addition, these non-Japanese fisheries are not subject to Japanese domestic fishery management and it is unclear what the catch and/or effort level of these non-Japanese fisheries would be in the future. This review concluded that these were not appropriate reasons to exclude catches from non-Japanese fisheries because there was enough information to develop catch-at-age time series using available data from a combination of Japanese and non-Japanese sources. Even if uncertain, it is inappropriate to assume the catch by these non-Japanese fisheries is zero, which is what is currently assumed. It is critical that all catches on the stock be included in an assessment model, especially VPA models, because the primary objective of an assessment model is to assess the impact of fishing (i.e, the input signal on the system) on the population dynamics of the stock (i.e., the response of the system to the signal). Neglecting these sources of catch would have biased the VPA model results, including estimates of spawning stock biomass (SSB), recruitment and fishing mortality (F). In some cases, it may be useful to develop model ensembles that cover a range of reasonable scenarios for these non-Japanese fisheries to help inform decision making. If the decision making

process requires focusing only on Japanese fisheries, it would be important to reconsider the current model-based assessment approach, and instead consider an observation-based management procedure that sets the stock's fisheries management (e.g., catch limits) based on a survey of the stock's biomass in Japanese waters.

The approach taken to develop the age compositions for the two sardine stocks (JS-P and JS-T) were generally similar and considered to be appropriate, albeit with room for improvement. However, the approach for the two jack mackerel stocks (JM-P and JM-T) were considered to be inappropriate and the resulting age compositions were not considered to be BSIA. For the sardine stocks, time- and area-specific age-length keys (ALKs) were developed from regular age readings of sampled hard parts for each month and year. The ALKs were subsequently used to convert sampled length compositions of the catch into age compositions. Importantly, such an approach allowed fish in a single length bin to be distributed into one or more ages, based on aging data. Annual length-weight relationships for different areas were also used to convert total catch in weight to total catch in numbers. In contrast, the jack mackerel stocks did not have regular age readings of sampled hard parts. Instead, the age composition data of the JM-P and JM-T stocks were based on cohort slicing of length composition data, which resulted in all fish from a single length bin being assigned to a single age. In addition, a single fixed growth curve was used for all areas (JM-P) and periods (JM-P and JM-T), which assumed that growth was constant for all areas and time. This assumption did not appear to be valid for these stocks, with unpublished data indicating that growth of these stocks varied substantially by area and over time. The modes of the length compositions appeared to vary over time and the assumed ages did not consistently match the modes in the length compositions. It is strongly recommended that a robust sampling program for the age composition of the catch using hard parts (e.g., otoliths), similar to the sardine stocks, be initiated for the jack mackerel stocks.

A complex variety of relative abundance indices were used to tune the VPA models, ranging from large-scale surveys to the sum of catch (without effort) from a single port. Some of the indices used were found to be appropriate but others were poorly developed. Most of the indices require substantial work to improve them. One concern was the use of nominal indices from commercial fisheries for several of the assessments. However, it was reported that work is currently ongoing to standardize these indices in time for the next assessments. Another concern was that the development and use of indices did not appear to be well thought out for some assessments. Available indices appeared to have been thrown into the model without considering if the indices were conflicting or representative of the stock. It is recommended that future assessments focus on a single or a small number of well-designed surveys that cover the entire stock's range, which will be vastly superior to a large number of smaller, uncoordinated surveys. One recommended approach for all the assessments is to develop a survey using the acoustic-trawl method, which combines acoustics and trawl data to survey the age-specific biomass of one or more stocks. Even if such a survey only samples Japanese waters, it can still be used as the primary source of information for the abovementioned observation-based management procedure.

The biological information and parameters used for all the assessments were considered appropriate but with substantial room for improvement. The stock structures for all four stocks appeared to be appropriate, given the current available information. However, more work on the stock structure is recommended, especially for the two jack mackerel stocks. The natural mortality (M) parameters used for all four assessments were within the expected range for the species (albeit maybe a little on the low side) and can be considered to be appropriate. However, there is also a clear need to re-evaluate and improve the M for all the stocks using up-to-date scientific information.

Given that the primary data inputs for all four assessments were not considered to be BSIA, this review concluded that the results of the VPA models were also not BSIA. This review did not spend much time on the VPA methodology because it was clear that the data issues identified precluded the VPA results from being considered as BSIA. One concern was that the treatment of uncertainty in all four assessments were considered inadequate. The JM-P stock assessment used bootstrapping of model residuals to estimate the uncertainty of the VPA model. However, it did not appear that this uncertainty was propagated into other aspects of the assessment. The bootstrap procedure also did not consider uncertainties in the data and biological parameters. The other three assessments did not consider uncertainty in the VPA model at all. Importantly, all four assessments assumed that the estimated population structure (N-at-age) and F-at-age in the terminal years of the VPA models were known without error, and these were in turn used to estimate the stock-recruitment relationships (SRRs) and in the projections. Given that the estimated N-at-age and F-at-age have uncertainties, the probability distributions from the projections will likely be erroneous if these uncertainties were not included. Along with the N-at-age and F-at-age, the uncertainty in the estimated recruitment and spawning biomass also appeared to be neglected. It is understood that working with uncertainties in VPA models is more difficult, especially when the data were not developed according to the assumptions made. However, incorporating and propagating uncertainty from the data and estimation models through to the projections is important because the probability distributions from the projections appeared to be important for decision making. If the projection results are critical, it may be beneficial to explore the use of statistical catch-at-age models, which can better estimate and propagate the uncertainties in the data and biological processes throughout the model.

All four assessments used the VPA model outputs to estimate the SRRs and in the short- and long-term projections. Given that the results from the VPA models were not considered to be BSIA, the short- and long-term projections were also considered to be not BSIA. All four assessments required short-term projections (2 years: JS-P, JM-P, and JM-T; 3 years: JS-T) to bring forward the estimated N-at-age in the terminal year of the VPA model to the current fishing year (2021) and set the allowable catch based on the projected SSB in 2021. In terms of methodology, the projection models were relatively straightforward and appropriate, given the information available. However, it is important to note that the projected SSB in 2021, on which the allowable catch is calculated from, is highly uncertain and the uncertainties in these projections are currently under-estimated. Basing the allowable catch on the short-term projections may lead to large mismatches between expectations and observations. One of the management objectives for the stocks was the probability of exceeding the target reference point (MSY) over the next 10 years should not exceed 50%. In addition, there are management objectives and probabilities associated with the limit and ban reference points. Long-term projections out to 2041 were used to estimate the appropriate F-multiplier (e.g., relative to F_{MSY}) to achieve the management objectives. The models used for the long-term projections were similar to the short-term projections, and were relatively straightforward and appropriate. However, similar to the short-term projections, the uncertainties for the long-term projections are currently under-estimated.

2. Background

Assessments of two Japanese sardine (*Sardinops sagax melanostictus*) and two Japanese jack mackerel (*Trachurus japonicus*) stocks were conducted by the Japan Fishery Research and Education Agency (FRA) in 2021. These four stock assessments were on the: 1) Japanese sardine - Pacific (JS-P) ; 2) Japanese sardine - Tsushima (JS-T); 3) Japanese jack mackerel - Pacific (JM-P; and 4) Japanese jack mackerel - Tsushima (JM-T) stocks. The assessments of the JS-P, JS-T, and JM-P stocks were led by the Fisheries Stock Assessment Center, Fisheries Resources Institute of the FRA, in collaboration with other scientific organizations. The JM-T stock assessment was jointly led by the FRA's Fisheries Stock Assessment Center, Fisheries Resources Institute; and the Environment and Application Department, Fisheries Technology Research Institute, in collaboration with other scientific organizations.

All four assessments were broadly similar and consisted of five phases. First, time series of catch-at-age and abundance indices were developed for each stock and a virtual population analysis (VPA) model was used to estimate the historical population dynamics. Importantly, the VPA model estimated the recruitment, spawning stock biomass (SSB), population structure (N-at-age), and fishing mortality at age (F-at-age) over the historical period. Second, the estimated recruitment and spawning biomass were used to develop the stock-recruitment relationship (SRR) and biological reference points for each stock. Third, short-term (two or three year) projections were used to bring forward the estimated N-at-age in the terminal year of the VPA model (e.g., 2019) to the current year (e.g., 2021). This short-term projection was performed using the abovementioned SRR and the estimated F-at-age from the average of several terminal years of the VPA model. Fourth, the projected SSB was used to determine the current stock status and, in combination with the harvest control rules (HCRs), used to determine the allowable biological catch (ABC). Lastly, long-term (e.g., 10 – 20 year) projections were used to test the robustness of the HCRs to reach management objectives.

3. Review System

An independent peer review of these assessments was organized by the Secretariat of Peer Review of the FRA under the specified Terms of Reference (ToR, Appendix 1). The review panel consisted of two reviewers from Japan, who were appointed by the Japanese Society of Fisheries Science, and myself as the overseas reviewer (Dr. Steven L. H. Teo, National Oceanic and Atmospheric Administration of the

USA). The ongoing COVID-19 pandemic and travel restrictions prevented a face-to-face review panel meeting. Instead, a short web meeting was held during September 13 and 14, 2021 for approximately four hours each day, and included English-Japanese translators to ease language difficulties. A separate meeting was held for the Japanese reviewers.

Assessment documents were translated from Japanese to English, and submitted before the web meeting. In order to help FRA scientists prepare for the web meeting, initial questions on the assessments were submitted prior to the web meeting, with follow up questions during the web meeting. The documentation for each assessment consisted of a main assessment document, several appendices, a document detailing the catch-at-age data, and a meeting report for determining reference points for each stock.

This report is an independent review of the four assessments based on my personal opinions of the submitted documents and clarifications during the web meeting. There was no communication between the Japanese reviewers and I. The submitted assessment documents were not detailed enough to conduct a thorough technical 'desktop' review. The lack of detailed documentation and time, combined with the online format and language difficulties, made it difficult to thoroughly review all aspects of the assessments. Nevertheless, this review attempts to cover the specified ToRs.

A presentation was also made at the start of the web meeting to explain the role of stock assessments in the management of domestic fisheries in Japan. It was noted that the results from these four stock assessments have already been used to determine Japanese domestic management for these stocks. Therefore, there was no opportunity to review and recommend alternative model configurations for the base case models. Instead, the primary aims for this review were to evaluate if the current stock assessments were based on the best scientific information available (BSIA) and recommend improvements for future stock assessments.

This report is subdivided into separate sections for each stock and a section summarizing the overall findings. As specified in the ToRs, the data, biology (i.e., information and parameters), methodology, and results from the assessment and projection models of each stock were reviewed and evaluated. The assessments of Japanese sardine stocks (JS-P and JS-T) were reviewed first, followed by the jack mackerel stocks (JM-P and JM-T). If a comment was applicable to multiple stocks, the reader will be referred to earlier sections whenever possible to reduce the report length and repetitiveness. However, the recommendations for each stock were kept separate for the sake of clarity.

4. Japanese Sardine – Pacific Stock (JS-P Stock)

4.1 Data

A VPA model was used to estimate the population dynamics of the JS-P stock during the historical period. A VPA model assumes that the catch-at-age is known without error. Therefore, it was critical to examine the data preparation and determine if this assumption was severely violated. In addition, the relative abundance indices used as tuning indices for the VPA model were also examined.

4.1.1 Catch-at-age

Developing the catch-at-age time series depends largely on two components: 1) the total catch in numbers; and 2) robustly sampling the age composition of the catch. Much of the discussion during the web meeting was focused on this part of the assessment. Overall, this review found that the catch-at-age data used for the JS-P stock assessment was not the BSIA because substantial amounts of total catch of this stock was available but not included in the assessment. Therefore, the assumption that the catch-at-age was known without error was severely violated, and likely led to large errors and biased the results of the VPA model and assessment.

The time series of total catch used in this assessment did not include catch by important non-Japanese fisheries fishing on the same stock. The total catch in this assessment was based on reported catches of the JS-P stock by Japanese fisheries. However, there were substantial catches by vessels from China and Russia that were excluded. For example, the North Pacific Fisheries Commission (NPFC) reported that the catches by China and Russia in 2019 were 133,000 and 46,000 t, respectively. It was also reported that there were no other non-Japanese sources of catch on this stock. This amount of catch by China and Russia fisheries is non-negligible relative to the catch by Japanese vessels of 521,000 t in 2019. Although the catches by China and Russia were reported to be important only in recent years, coinciding with an expansion in range by the stock, neglecting these sources of catch would have resulted in biased results, including estimates of SSB, recruitment and fishing mortality (F), in recent years.

A model run including the catch by non-Japanese fisheries was conducted (Model A9; see Appendix 9 in the assessment document), with the age composition for the China and Russia fisheries assumed to be the same as the Japan purse seine fleet operating in Northeastern Japan. The results from Model A9 were examined during the web meeting and compared with the base case model in the stock assessment. This comparison indicated important differences in the results of the assessment (Table 1). Importantly, the estimated %SPR in 2019 for Model A9 was visibly lower (indicating higher F) than the base case model (Fig. 1) and F2019/FMSY was also >1, which indicated overfishing relative to MSY was occurring, in contrast to the base case model.

Table 1. Comparison of selected results for 2019 of the base case model (Japan catch only) and Model A9 (includes China and Russia catches).

	Base Case Model	**Model A9**
Total catch 2019	521,000 t	700,000 t
SSB 2019	1.59 Mt	1.72 Mt

	Base Case Model	Model A9
Recruitment (2017-2019)	44.1 M	48.3 M
SSB2019/SSB_MSY	>1	>1
F2019/F_MSY	<1	>1

There was substantial discussion on the decision to not include catches from China and Russia in the base case model, even though the catch was reported to the NPFC. One reason was that the catches from China and Russia only increased to high levels in recent years as the stock's abundance increased and the range expanded outside the Japanese exclusive economic zone (EEZ). Previous to 2016, there were only limited catches outside the Japanese EEZ. Therefore, previous assessments have only included catches from Japanese fisheries. Second, the age distribution of the catch from these non-Japanese fisheries is unknown. Third, it was unclear if the high non-Japanese catches would continue in the future, and if so, what those levels would be. Fourth, the vessels from China and Russia fishing in the high seas would not be subject to Japanese domestic management. Therefore, the decision was made to not include these catches into the base case VPA model and future projections. Based on this discussion, this review concluded that these were not appropriate reasons to exclude catches from non-Japanese fisheries. It is critical that all catches on the stock be included in the VPA model, even if previous assessments have not done so. If age distribution of these fisheries are not reported, it was recommended to estimate the catch-at-age based on similar Japanese fisheries, as was done for Model A9. The recommendations on incorporating catches from non-Japanese fisheries into future projections would be discussed in Section 4.4.4.

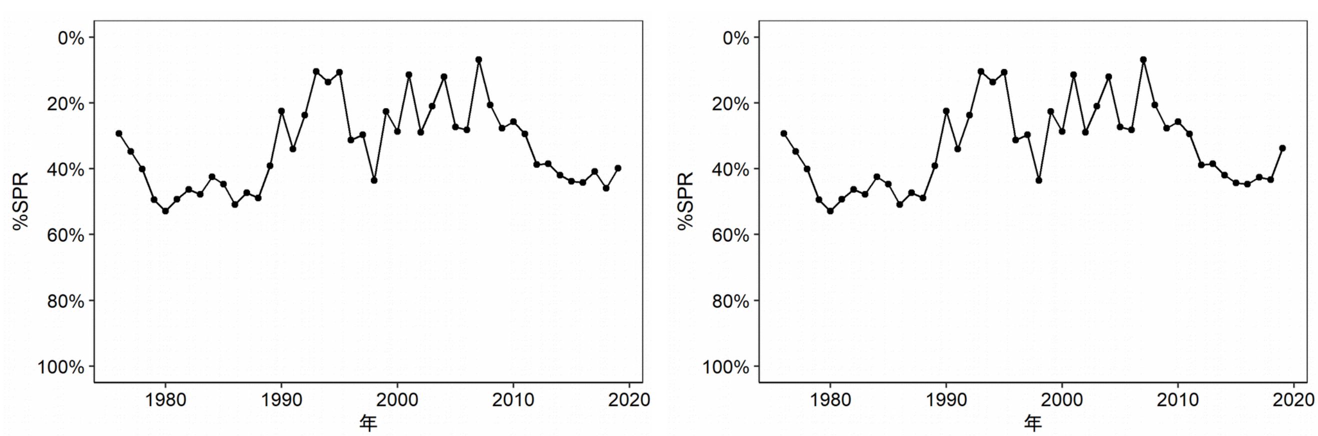

Figure 1. Comparison of the %SPR time series of the base case model (Japan catch only; left panel) and Model A9 (includes China and Russia catches; right panel). Lower values of %SPR indicate higher fishing intensity and F. Note that lower %SPR values are towards the upper part of the y-axis.

Age-length keys (ALKs) were developed for two areas (Northeast and Southwest Japan) for each month and year to convert length distributions into age distributions. Annual length-weight relationships for each prefecture were used to convert total catch in weight to total catch in numbers. The total catch in numbers were combined with the age distributions to obtain the catch-at-age data used in the VPA model. The ALKs were based on length measurements and age readings of scales from sampled sardines. There was no discussion of the sampling designs for the length distributions and scale samples due to lack of time. It is assumed here that the sampling designs were robust and appropriate but should be reviewed in the future. It was noted that reading scales are more likely to lead to biased ages as compared to reading otoliths. Therefore, it would be useful to perform experiments to compare different hard parts for aging and to validate the ages. In general, the approach taken for the ALKs was appropriate. However, this review noted that age readings were predominantly made by single readers and aging errors were not estimated on an operational basis. If validation experiments are not feasible, it would nevertheless be useful to obtain reader-specific estimates of bias and error (Punt et al. 2008). In addition, the age-length keys had some bins with missing data, likely due to insufficient sampling, which may cause some bias and error in the age compositions. If increasing sampling is not possible, it may be useful to experiment with interpolating those gaps with fitted seasonal growth models. Although cost-prohibitive, operational sampling and aging of the catch is commonly thought to result in improved age compositions compared to using age-length keys. Finally, this review also recommends that a study of the variability in the length-weight relationships between months or seasons for different areas be conducted to determine if it is important to have finer temporal scale length-weight relationships.

4.1.2 Relative abundance indices

Three tuning indices were used in the JS-P assessment: 1) an egg production index (EPI) was used to represent trends in the SSB; 2) the Northwestern Pacific index (NWPI) is an age-0 index based on an acoustic pelagic fish survey; and 3) the wintering juveniles index (WJI) is an index for age-1 sardines based on the CPUE of commercial large-scale purse seine fishery. However, based on the discussion during the web meeting, this review does not consider the tuning indices to be BSIA.

The EPI is based on a large scale monthly egg and larval survey conducted by 19 prefectural fisheries institutes and the FRA. The egg and larval survey extends from the southern part of the East China Sea to the north of Tohoku, and covers the spatial extent of the spawning ground during both low and high abundance periods. The acoustic pelagic fish survey for the NWPI was conducted by the FRA during September and October, using a midwater trawl and quantitative echo sounder, and covered a broad area from Hokkaido to Joban. The WJI is the cumulative CPUE of small and medium sized fish (approximately age-1 fish) caught by the large-scale purse seine fishery from December of the previous year to April of the current year.

There were limited details in the documentation and not enough time during the web meeting to review the indices in detail. Nevertheless, it was noted that all three indices were nominal indices (i.e., unstandardized), with an implicit assumption that catchability was constant for these indices. Examining potential catchability changes in an abundance index and correcting for these changes (i.e., standardizing the index) is important, especially for indices not based on surveys, before fitting an index in a stock assessment (Maunder and Punt 2004). It was reported during the web meeting that there are ongoing attempts to standardize these indices and standardized indices will likely be used in future assessments. However, only nominal indices were used for this assessment and there was no evidence to support the assumption of constant catchability. Therefore, the abundance indices in this assessment are likely a function of both abundance and catchability, and this review does not consider the abundance indices to be BSIA.

It is strongly recommended that future assessments examine the catchability of these indices and use standardized indices, where appropriate. It is also recommended to estimate the uncertainty of the indices. Data from the egg and larval, and acoustic pelagic fish surveys are used to develop indices for a variety of species. It is therefore recommended that dedicated reviews of these surveys be used to examine the survey design in detail. It is also recommended that future assessments focus on the EPI and NWPI because the WJI is more an index of school size integrated over the fishing season rather than population abundance. In addition, there are also potential improvements for the EPI and NWPI that may be investigated. Currently, the EPI is an aggregate of the eggs sampled per unit effort of the surveys. However, the egg production of individual fish is affected by various biological parameters like batch fecundity, spawning fraction, sex ratio, and average fish weight. If adult fish are sampled in conjunction with the egg samples, the daily egg production method may be used to estimate SSB from the egg surveys (Lo et al. 2005). If the biological sampling required for the daily egg production method is not available, it would be useful to consider a spatiotemporal model for standardization. The NWPI can also be improved by combining the acoustics and midwater trawl data (i.e., acoustic-trawl method) to obtain length- or age-specific abundance (Zwolinski et al. 2016).

4.1.3 Recommendations on data

Based on the above findings, there are several recommendations for improving the data used for the JS-P assessment:

1) It is critical to include catches from all fisheries, including non-Japanese fisheries, that capture this stock of sardine for the assessment.

2) It is recommended to perform experiments to compare different hard parts for aging and to validate the aging.

3) It would also be recommended to obtain reader-specific estimates of bias and error in the aging.

4) If possible, it would be useful to increase sampling to minimize the gaps in the ALKs. Otherwise, it would be recommended to experiment with interpolating those gaps with fitted seasonal growth models.

5) It is recommended to conduct a study to compare the length-weight relationships between months or seasons for different areas to determine if it is important to have finer temporal scale length-weight relationships.

6) It was not possible to review the robustness of the length and age sampling program but it would be recommended to review it in the future.

7) It is strongly recommended that future assessments examine the catchability of these indices and use standardized indices, where appropriate.

8) It is also recommended to estimate the uncertainty of the indices.

9) It is recommended that dedicated reviews be used to examine the survey design of the egg and larval fish survey, and the pelagic fish survey in detail.

10) It is recommended that future assessments focus on the EPI and NWPI because the WJI is more an index of school size integrated over the fishing season rather than population abundance.

11) If possible, adult fish should be sampled in conjunction with the egg samples, so that the daily egg production method may be used to improve SSB estimates from the egg surveys.

12) It is strongly recommended to investigate the acoustic-trawl method for estimates of SSB and/or length- and age-specific abundance from the acoustic pelagic fish survey data.

4.2 Biology

4.2.1 Stock structure and distribution

The stock structure for this assessment was appropriate, given the current available information. However, it is encouraged to continue research on this topic. There are two groups of fish: 1) offshore, and 2) coastal, within this stock. Both groups are thought to be from the same spawning grounds in the Kuroshio Current but the proportion of each cohort that goes to each group is highly variable and depends on oceanic conditions, with certain months favoring one group or the other.

4.2.2 Natural mortality

The natural mortality (M) of the JS-P stock was assumed to be 0.4 y^{-1} based on a reported maximum age (A_{max}) of 7 years, using the relationship of $M=2.5/A_{max}$ from Tanaka (1960). However, a recent reanalysis of the $M=a/A_{max}$ relationship (i.e., a is the fitted parameter) with substantially more data, suggested that a is ~5 (Then and Hoenig 2014; Then et al. 2015), which would in turn suggest a higher M for this stock. In addition, the commonly-used empirical relationship by (Hoenig 1983) that relates maximum age with M would also suggest a higher M for this stock. In comparison, a recent assessment of sardine in the Eastern Pacific Ocean (Kuriyama et al. 2020) had an estimated posterior of 0.585 y^{-1} for M. However, it should be noted that prior to 2017, the sardine assessment in the Eastern Pacific Ocean had also assumed a fixed M of 0.4 y^{-1}.

These lines of evidence suggest that the M used for this assessment is within the expected ballpark for this species (albeit maybe a little on the low side) and can be considered to be appropriate, given current understanding. However, there is also enough evidence to suggest that the M needs to be re-evaluated and improved. Given the important of M in stock assessments (Punt et al. 2021), It is therefore strongly recommended that research on M be conducted prior to the next assessment. It would be recommended to perform a meta-analysis (Hamel 2015) using more recent empirical relationships and with more data than Tanaka (1960). In addition, it would be recommended to obtain the maximum age from otolith samples because ages from scale samples tend to biased low, especially for higher ages.

4.2.3 Maturity

The maturity of the JS-P stock exhibited substantial variability between areas and years. In some years and areas, a larger proportion of age-1 fish appeared to be mature. In this assessment, it is assumed that the maturity schedule for 2016 – present is the same as for 1994 – 1997. Given the available data, this is assumption appears to be appropriate. However, the effect of this assumption on assessment results is unclear. It is recommended that research continue to be conducted on the maturity of this stock: 1) continue data collection on maturity by area and year; and 2) examine the effects of this maturity variability on the population dynamics and assessments of this stock.

4.2.4 Growth

The sampling program for this stock included sampling of hard parts for aging. Therefore, growth was not important for this assessment. Given that, there was no review of the growth for this stock. Nevertheless, it is recommended to include growth studies in the future to help with interpolating data gaps in the ALKs (see Section 4.1.3).

4.2.5 Recommendations on biology

Based on the above findings, there are several recommendations for improving the biological assumptions and parameters used for the JS-P assessment:

1) It is strongly recommended that the natural mortality for this stock be re-examined by performing a meta-analysis using empirical relationships that are more recent and richer in data.

2) It is recommended to estimate the maximum age from otolith samples rather than scale samples because ages from scale samples tend to biased low, especially for higher ages.

3) It is recommended that data on maturity by area and year continue to be collected and the effects on population dynamics be examined.

4.3 Estimation Model

In Section 4.1, this review found that the primary data inputs (i.e., catch-at-age and abundance indices) for the assessment of the JS-P stock were not BSIA. Given that the estimation model was a VPA, which makes the critical assumption that the catch-at-age is known without error, this review also found that the results of the VPA were not BSIA. This review did not spend much time on the VPA methodology during the web meeting because it was clear that the data issues identified precluded the VPA results from being considered as BSIA. Nevertheless, this review briefly examined the methodology of the VPA model.

4.3.1 VPA model

A modified ridge VPA model (Okamura et al. 2017) was used in the JS-P stock assessment to estimate the historical population dynamics of the stock. The ridge VPA reduces the retrospective bias in model results by including a penalty term for fishing mortality (F) in the terminal years that is weighted by the size of the retrospective bias. A weighting coefficient (λ, between 0 and 1) sets the relative weighting between the fit to the indices ($\lambda \rightarrow 0$) and the F penalty ($\lambda \rightarrow 1$). As λ approaches 0, the model is similar to a traditional tuned VPA, but as λ approaches 1, the model prioritizes reducing the F penalty over model fit. The λ is estimated from minimizing the retrospective bias of biomass over a number of terminal years. Essentially, there is a tradeoff between reducing the retrospective bias (and terminal F penalty) versus improving model fit to the tuning indices. This assessment modified the ridge VPA by including an additional weighting coefficient (η, between 0 and 1) that sets the relative weight of F for age-0 fish in relation to the Fs for age-1+ fish in the F penalty. This η weighting coefficient was included to minimize an observed tradeoff between lowering the bias in spawning biomass estimates and a higher negative bias in recruitment. As η approaches 1, the F penalty term becomes increasingly based on F at age-0 at the expense of the Fs for age-1+.

A brute force search (intervals of 0.01) was used to set the λ and η for the base case model, with the objective of minimizing the retrospective biases of various population estimates. The λ and η were set at 0.93 and 0.99, respectively, which are extremely high weighting values towards the F penalty term and F at age-0 in particular. In effect, the estimates from the VPA model becomes overwhelmingly based on the F at age-0 at the expense of fit to the indices and Fs at age-1+. For example, the weighting of the F at age-0 is 99-fold greater than for ages 1 to 5+. Minimizing the retrospective is a good objective but such extreme weights do not appear to be reasonable. It is recommended that simulation experiments be conducted on this modified ridge VPA to show that such extreme weights are reasonable and that the estimated population dynamics are not heavily biased.

The necessity of such extreme weights may be an indication of fundamental problems with the data (e.g., missing catch, unrepresentative indices) and biological parameters (e.g., natural mortality). Simply changing the model settings likely only alleviated the symptoms for this assessment but did not solve the fundamental problems with the assessment. A VPA makes strict assumptions about the data, which have been severely violated in this assessment (Section 4.1). The priority would be to solve the problems with the data first. Although improving the data for this assessment is critical, it is also recommended to explore the use of statistical catch-at-age models, which can be more flexible and able to model the data as it is collected. It would also be easier to develop models with alternative hypotheses and examine conflicts in the data and parameters.

4.3.2 Model diagnostics

The assessment documents did not provide model diagnostics like retrospective patterns and model fit. However, these model diagnostics for the base case model were shown during the web meeting. The model diagnostics indicated that the SSB index was not well fit, especially over the last few years when the SSB increased. There were negligible retrospective patterns, which was not surprising given the use of the modified ridge VPA (see Section 4.3.1).

4.3.3 Uncertainty

The treatment of uncertainty in the JS-P assessment was inadequate. Non-parametric bootstrapping of the residuals of the abundance indices were used to generate 10,000 bootstrapped indices to represent the uncertainty in the model fit to the indices and used in iterations of the VPA to generate 90% confidence intervals. This resulted in large uncertainties in the terminal year estimates from the VPA model. However, it did not appear that this uncertainty was propagated into other aspects of the assessment. The bootstrap procedure also did not consider uncertainties in the data and biological parameters. Importantly, the estimated population structure (N-at-age) and F-at-age in the terminal years were assumed to be known without error to estimate the stock-recruitment relationship (SRR) and projections. Given that the estimated N-at-age and F-at-age have uncertainties that were unaccounted for, the estimated probability distributions from the projections will likely be erroneous. Along with the N-at-age and F-at-age, the uncertainty in the estimated recruitment and spawning biomass also appeared to be neglected. It is understood that it is more difficult to work with uncertainty in a VPA model, especially when the data were not developed according to the assumptions made (e.g., catch-at-age is known without error). It may be beneficial to explore the use of statistical catch-at-age models, which can more easily include the uncertainty in the data and biological processes throughout the model.

4.3.4 Recommendations on estimation modeling

Based on the above findings, there are several recommendations for improving the estimation models used for the JS-P assessment:

1) It is strongly recommended that model diagnostics be included in the assessment documents.

2) It is recommended that simulation experiments be conducted on this modified ridge VPA to show that the extreme weights used in this assessment are reasonable and that the estimated population dynamics are not heavily biased.

3) It is critical that the uncertainty be treated appropriately in the stock assessment. It is therefore strongly recommended that future assessments do not assume that estimated quantities (e.g., N-at-age, F-at-age, recruitment, spawning biomass) are known without error and, where appropriate, uncertainty in the data collection and biological processes be included and propagated through the estimation model.

4) It would be beneficial to explore the use of statistical catch-at-age models, which can be more flexible and more easily include the uncertainty in the data and biological processes throughout the model. It would also be easier to develop models with alternative hypotheses.

4.4 Projections

The outputs from the VPA model were used to estimate the stock-recruitment relationship (SRR) and used in the short- and long-term projections. Given that the results from the VPA model were not considered to be BSIA (Section 4.3), this review also concluded that the results of the projections were not BSIA. Nevertheless, this review briefly examined the methodology of the projections.

4.4.1 Stock-recruitment relationship (SRR)

The estimated recruitment and SSB from the VPA model was used to develop the appropriate SRR to use for calculating biological reference points (BRPs) and for the future projections. It was reported during the web meeting that the assessment scientists follow guidelines from FRA to estimate the SRR in a stock assessment. These guidelines appear to be used across all FRA stock assessments and is beyond the ToRs for this review. Therefore, this section does not consider whether the SRR was BSIA.

Regime shifts are important to the population dynamics of sardines and has been documented for this stock. An important discussion during the web meeting was how to identify a regime shift in the JS-P stock and which regime to use to estimate stock status and in the projections. It was clear that the recruitment for the JS-P stock has increased substantially over the past few years. However, the evidence was mixed on whether this was the start of a high recruitment regime or just variability within the current regime. Therefore, this review thought that the use of the current 'normal' regime was reasonable. However, it is recommended that the criteria for classifying the regime be explained clearly in the assessment documents. It is also recommended that the recruitment and SSB estimates used to develop the SRRs have uncertainty associated with them.

4.4.2 Short-term projections

The terminal year of the estimation model is 2019 but the allowable catch for the current fishing year is based on SSB in 2021. Therefore two years of projections were required to bring forward the estimated N-at-age in 2019 to the projected spawning biomass in 2021. Given that the results from the VPA model were not considered to be BSIA, it follows that the results of these projections were also not considered to be BSIA because the projections relied on the VPA model results for inputs.

In terms of the methodology, the projection models are relatively straightforward and appropriate, given the information available. However, the problems with using short-term projections to estimate SSB in 2021 and set allowable catch should be noted. One of the problems is that the projection results are highly sensitive to the estimates of age-0 and age-1 fish in the terminal year of the VPA model but these estimates are highly uncertain due to the lack of observations. In addition, catches by non-Japanese fisheries are not limited by Japanese management. Therefore, the projected SSB in 2021, on which the allowable catch is calculated from, is highly uncertain and may be biased. It is recommended that Japanese scientists work with managers to develop ways to incorporate non-Japanese fisheries into these projections. One possibility is to develop multiple scenarios for non-Japanese fisheries, and provide managers with the results from all the scenarios. It should also be noted that the uncertainties in these projections are currently under-estimated because the N-at-age in the terminal year of the estimation model is assumed to be known without error, which is considered inappropriate. It is therefore important and recommended to incorporate that uncertainty in the N-at-age in the terminal year into the short term projections. The uncertainty in the projected recruitment is also likely under-estimated.

Basing the allowable catch on the two-year projections of spawning biomass may also lead to large mismatches between expectations and observations. Such mismatches have led other countries to use management procedures based on observations from surveys before the start of the fishing season. For example, this management procedure could set the allowable catch based on a survey of the stock's biomass before the fishing season starts. This survey would alleviate the problems of knowing and managing a stock with highly variable recruitment driven largely by environmental conditions, as well as the catch by non-Japanese fisheries. However, it would be important to use simulations (e.g., an MSE-type process) to evaluate if such a procedure would meet management objectives and be robust for this stock, including both Japanese and non-Japanese fisheries that are not under the management procedure.

4.4.3 Long-term projections

One of the management objectives for the JS-P stock was the probability of exceeding the target reference point (MSY) over the next 10 years should not exceed 50%. In addition, there are management objectives and probabilities associated with the limit and ban reference points. Long-term projections out to 2041 were used to estimate the appropriate F-multiplier (e.g., relative to F_{MSY}) to achieve the management objectives. Similar to the short-term projections, the results of these long-term projections were not considered to be BSIA because the inputs for the projections relied on the VPA model results, which were not considered to be BSIA.

The models used for the long-term projections were similar to the short-term projections, and were relatively straightforward and appropriate. However, the uncertainty for the long-term projections are currently under-estimated because the N-at-age in the terminal year of the estimation model is assumed to be known without error, which is considered inappropriate (Section 4.3.3). It is therefore important to incorporate that uncertainty in the N-at-age in the terminal year into both the short term and long-term projections. This would be especially important for calculating the probabilities of exceeding the limit and ban reference points, which are based on the tail of the probabilities. In addition, catches by non-Japanese fisheries are not limited by Japanese management. Therefore, it may not be reasonable to assume that the F-multiplier applies to non-Japanese fisheries as well. Similar to the short-term projections, one possibility may be to develop multiple scenarios for non-Japanese fisheries, and provide managers with the results from all the scenarios.

Another important consideration is that of potential changes in recruitment regime over the projection period. An alternative recruitment scenario incorporating some elements of this was developed for use in projections incorporating the recent improved recruitment trend (Appendix 8 of the assessment document). This alternative recruitment scenario was based on having a resampling procedure that only samples the last 5 years of the assessment period (2015 – 2019) for the first 5 years of projections (2022 -2026), and expands the resampling period by 5 years (2010 - 2019) for the next 5 years of projections (2027 – 2031), and so on until 2041. This procedure has the same effect as assuming some autocorrelation in recruitment. However, the procedure is relatively arbitrary and may not be statistically sound. One recommendation would be to estimate the statistical distribution (potentially bimodal) and the autocorrelation of the estimated recruitment deviates, and use those statistics to represent future recruitment. This allows the projected recruitment to be based on the statistical distribution of the previously observed regimes.

4.4.4 Recommendations on projections

Based on the above findings, there are several recommendations for improving the SRRs and projections used for the JS-P assessment:

1) It is recommended that the criteria for classifying the regime be described clearly in the assessment documents so that the criteria can be consistently used in future assessments.

2) It is recommended that the recruitment and SSB estimates used to develop the SRRs have uncertainty associated with them.

3) It is recommended to incorporate that uncertainty in the N-at-age in the terminal year into the short term and long-term projections.

4) It is recommended to develop multiple scenarios for the projected catch of non-Japanese fisheries, and provide results from all the scenarios.

5) It is recommended to estimate the statistical distribution (potentially bimodal) and autocorrelation of the estimated recruitment deviates, and use those statistics to represent future recruitment in the long-term projections.

5. Japanese Sardine – Tsushima Stock (JS-T)

The stock assessment of the JS-T stock was closely related to the JS-P stock. Therefore this review found similar problems and provided similar recommendations. In order to reduce the repetitiveness and lengthiness of the report, references are made to earlier sections when appropriate. However, the recommendations for the JS-T stock were kept separate for the sake of clarity.

5.1 Data

A VPA model was used to estimate the population dynamics of the JS-T stock during the historical period. Therefore, the assessment of the JS-T stock made similar assumptions on data as the JS-P stock (see Section 4.1). One key difference between the JS-T and JS-P data was that the JS-T catch was considered to be abnormally low in 2019 due to distribution shifts, and was not used in the assessment. Therefore, the terminal year for the JS-T data in this assessment was 2018.

5.1.1 Catch-at-age

Similar to the JS-P stock, the JS-T stock assessment likely had substantial errors in the total catch and hence catch-at-age data (see Section 4.1.1 for details). The critical assumption for VPA models that the catch-at-age was known without error was severely violated and likely led to large errors and biased the results of the VPA model and assessment. Overall, this review found that the catch-at-age data used for the JS-T stock assessment was not BSIA because substantial amounts of total catch of this stock was available, albeit uncertain, but not included in the assessment.

Critically, the time series of total catch used in this assessment did not include large amounts of catch by important non-Japanese fisheries fishing on the same stock. The total catch in this assessment was based on reported catches of the JS-T stock by Japanese fisheries but did not include catches by vessels from China, Korea, and Russia. These catches by non-Japanese fisheries were at least in excess of 100,000 t annually. Similar to the JS-P assessment, even though Japanese fisheries were the main source of catch for this stock, neglecting these non-Japanese sources of catch would have resulted in biased results, including estimates of SSB, recruitment and fishing mortality (F). Although the catches by non-Japanese fisheries were uncertain, there appears to be enough information to develop a range of reasonable historical catch scenarios for the assessment. Therefore, it is critical that future assessments include the catches from non-Japanese fisheries.

The monthly ALKs for the JS-T stock assessment were developed in a similar manner to the JS-P stock, using scales sampled from the catch in two areas (i.e., East China Sea and the Sea of Japan). Similarly, the approach taken for the ALKs for the JS-T stock was appropriate, albeit with room for improvements. There were also similar recommendations on potential improvements for future stock assessments (Section 5.1.3).

5.1.2 Relative abundance indices

Three tuning indices were used in the JS-T assessment: 1) an egg production index (EPI) was used to represent SSB trends; 2) the Sakai index was used to represent the biomass of age-0+ (i.e., all ages) fish; and 3) the Ishikawa index represented the age-1+ biomass. Similar to the JS-P assessment, this review did not consider these tuning indices to be BSIA.

The most important index was the EPI, which represented the SSB relative abundance trends and was based on a large scale egg and larval survey during January - June. The survey extended from waters off Kagoshima Prefecture in western Kyushu to waters off Aomori Prefecture in the Sea of Japan, which corresponded to the known spawning areas of the stock. The Sakai and Ishikawa indices were based on CPUEs from purse seine vessels based in the Sakai and Ishikawa ports, respectively. For the Sakai index, individual vessel data were not available. Instead, the index was standardized from the daily port level CPUE calculated by the total amount of catch for the port divided by the number of vessels for the day. For the Ishikawa index, individual vessel logbook data were used to calculate the CPUE. However, the vessels were thought to target sardine for only a portion of the effort, and the directed CPUE approach (Biseau 1998) was used to identify effort targeted at sardines. However, the CPUEs were not standardized to correct for variables that may have affected catchability (e.g., seasons/months, vessel effects, area).

There were limited details in the documentation and not enough time during the web meeting to review the indices in detail. It was nevertheless noted by this review that only the Sakai index was standardized while the EPI and Ishikawa index were nominal indices. Although the Sakai index was standardized, the standardization was probably inadequate to correct for the use of port level data. Importantly, the Sakai and Ishikawa indices were based on purse seine fisheries and the CPUEs, as calculated, were more closely related to school size rather than population abundance. Based on the discussion during the web meeting, the operational areas for both fisheries were also relatively small compared to the stock range, and likely did not represent the stock as a whole. This could be seen in 2014 – 2015, when the distribution of the stock appeared to have changed and affected the catchability. Given the problems with these indices, it was recommended that these indices not be used in future assessments to represent the relative abundance trends for this stock. The EPI was promising, and would have been considered BSIA, if it was standardized. Most promisingly, the assessment documents also described an acoustic pelagic survey for the JS-T stock in the waters off Kyushu and the Sea of Japan, using a midwater trawl and quantitative echo sounder. This survey appears similar to the NWPI survey for the JS-P stock but was not used for this assessment because the time series was currently considered too short. For future assessments, it was recommended to develop an index by combining the acoustics and midwater trawl data (i.e., acoustic-trawl method) to obtain length- or age-specific abundance (Zwolinski et al. 2016). It was strongly recommended that future assessments focus on the EPI and the acoustic pelagic survey, and a review be conducted specifically focused on the design for these surveys. In addition, there were also potential improvements suggested for the EPI. Similar to the EPI for the JS-P stock, the EPI for the JS-T was also an aggregate of the eggs sampled per unit effort of the surveys. Therefore, it was also similarly recommended to use the daily egg production to estimate SSB from the egg surveys (see Section 5.1.2). If the biological sampling required for the daily egg production method is not available, it would be useful to consider a spatiotemporal model for standardization.

5.1.3 Recommendations on data

Based on the above findings, there are several recommendations for improving the data used for the JS-T assessment:

1) It is critical to include catches from all fisheries, including non-Japanese fisheries, that capture this stock of sardine for the assessment.

2) It is recommended to perform experiments to compare different hard parts for aging and to validate the aging.

3) It is recommended to obtain reader-specific estimates of bias and error in the aging.

4) If possible, it would be useful to increase sampling to minimize the gaps in the ALKs. Otherwise, it is recommended to experiment with interpolating those gaps with fitted seasonal growth models.

5) It is recommended to conduct a study to compare the length-weight relationships between months or seasons for different areas to determine if it is important to have finer temporal scale length-weight relationships.

6) It was not possible to review the robustness of the length and age sampling program but it is recommended to review it in the future.

7) It is strongly recommended that future assessments examine the catchability of these indices and use standardized indices, where appropriate.

8) It is also recommended to estimate the uncertainty of the indices.

9) It is recommended that dedicated reviews be used to examine the survey design of the egg and larval fish survey, and the pelagic fish survey in detail.

10) It is strongly recommended to develop an index from the acoustic pelagic survey by using the acoustic-trawl method to obtain estimates of SSB and/or length- and age-specific abundance from the acoustic pelagic fish survey data.

11) It is recommended that future assessments focus on the EPI and the index from the acoustic pelagic survey rather than the Sakai and Ishikawa indices because the latter indices are more of indicators of school size rather than population abundance, and are also based on fisheries with relatively small operational areas compared to the stock range.

12) If possible, adult fish should be sampled in conjunction with the egg samples, so that the daily egg production method may be used to improve SSB estimates from the egg surveys.

5.2 Biology

5.2.1 Stock structure and distribution

The stock structure for this stock assessment appeared to be appropriate, given the currently available information. However, it is noted that the JS-T and JS-P may be mixed in some areas and is recommended to continue research on this topic. It is also recommended to better understand the distribution of this stock, especially in non-Japanese waters, because changes in the distribution in some years have resulted in some years of data being unusable.

5.2.2 Natural mortality

The M of the JS-T stock was assumed to be 0.4 y^{-1}, using the relationship of $M=2.5/A_{max}$ from Tanaka (1960). For similar reasons to the JS-P assessment (Section 4.2.2), it was likely that the M used for this assessment was within the expected ballpark for this species (albeit maybe a little on the low side) and considered to be appropriate, given current understanding. However, there was also enough evidence to suggest that the M should be re-evaluated and improved. It was therefore strongly recommended that research on M be conducted prior to the next assessment. It would be recommended to perform a meta-analysis (Hamel 2015) using more recent empirical relationships and with more data than Tanaka (1960). In addition, it would be preferable to obtain the maximum age from otolith samples because ages from scale samples tend to biased low, especially for higher ages.

5.2.3 Maturity

The maturity of the JS-T stock exhibited substantial variability between areas and years. In some years and areas, a larger proportion of age-1 fish appeared to be mature, especially during periods of low abundance. In this assessment, it was assumed that the maturity schedule for 2016 – 2019 is the same as for other periods with similar stock status (e.g., 1960 – 1970, 1994 – 1999). Given the available data, this assumption appeared to be appropriate. However, the effect of this assumption on assessment results was unclear. It was recommended that research continue to be conducted on the maturity of this stock: 1) continue data collection on maturity by area and year; and 2) examine the effects of this maturity variability on the population dynamics and assessments of this stock.

5.2.4 Growth

The sampling program for this stock included sampling of hard parts for aging. Therefore, growth was not important for this assessment. Given that, there was no review of the growth for this stock. Nevertheless, it was recommended to include growth studies in the future to help with interpolating data gaps in the ALKs (see Section 4.1.3).

5.2.5 Recommendations on biology

Based on the above findings, there were several recommendations for improving the biological assumptions and parameters used for the JS-T assessment:

1) It is strongly recommended that the natural mortality for this stock be re-examined by performing a meta-analysis using empirical relationships that are more recent and richer in data.

2) It is recommended to estimate the maximum age from otolith samples rather than scale samples because ages from scale samples tend to biased low, especially for higher ages.

3) It is recommended that data on maturity by area and year continue to be collected and the effects on population dynamics be examined.

4) It is recommended to continue research on the stock structure and distribution of this stock.

5.3 Estimation Model

Similar to the JS-P stock, this review found that the results of the VPA model of the JS-T stock were not BSIA because the primary data inputs were not BSIA. This review did not spend much time on the VPA methodology during the web meeting because it was clear that the data issues identified precluded the VPA results from being considered as BSIA.

5.3.1 VPA model

A standard tuning VPA model was used in the JS-T stock assessment to estimate the historical population dynamics of the stock. It was noted that the VPA model's terminal year was 2018 instead of the expected 2019 because the catches in 2019 were very low due to changes in distribution, and using 2019 as the terminal year would have resulted in unreasonable estimates from the VPA model. The F_{2019} and SSB_{2019} were projected from the 2018 model estimates, and used to determine stock status in 2019. There may have been legal and management considerations in taking such an approach, and is therefore beyond the scope of this review. However, this review would note that taking such an approach likely resulted in highly uncertain F_{2019} and SSB_{2019} being used for stock status determination, and the large uncertainties were likely vastly underestimated and not adequately included in the stock status determination and assessment.

This review noted that the strict assumptions about the data for a VPA model were severely violated for this assessment, which was similar to the JS-P stock (see Section 4.1). The priority for future improvements would therefore be to solve the problems with the data first. It was also recommended to explore the use of statistical catch-at-age models, which can be more flexible and able to model the data as it is collected. However, given the difficulties of obtaining data from non-Japanese fisheries for this stock, it may not be possible to improve the data in the near future. If the JS-T catch of non-Japanese fisheries remains poorly known in the future, it may be more appropriate to use alternative assessment and management methods for this stock. For example, an observation-based management procedure that sets the catch limits based on a survey of the stock biomass in Japanese waters before the fishing season starts, could be used to manage the stock for Japanese fisheries. This would also reduce the problems with predicting the recruitment regime for future projections. However, it would be important to use simulations (e.g., an MSE-type process) to evaluate if such a procedure would meet management objectives and be robust for this stock, including both Japanese and non-Japanese fisheries that are not under the management procedure. Alternatively, data-poor methods can be used to assess the stock but the associated uncertainty would be relatively high and may not be appropriate for the management of this stock.

5.3.2 Model diagnostics

The assessment documents did not provide model diagnostics like retrospective patterns and model fit. However, these model diagnostics for the base case model were shown during the web meeting. The model diagnostics indicated that the indices were poorly fit. There were some retrospective patterns but these patterns were not severe.

5.3.3 Uncertainty

The treatment of uncertainty in the JS-T assessment was inadequate. As discussed in the above sections, there were large uncertainties throughout the assessment. However, there was minimal consideration of uncertainty in the data, modeling, and results of the assessment. For example, the estimated population structure (N-at-age) and F-at-age in the terminal years were assumed to be known without error. This was an assumption that was severely violated for this assessment and not recommended for any assessment. Therefore, when the N-at-age and F-at-age were used in the projections without uncertainty, the estimated probability distributions from the projections were likely be erroneous. Along with the N-at-age and F-at-age, the uncertainty in the estimated recruitment and spawning biomass also appeared to be neglected. It is understood that it is more difficult to work with uncertainty in a VPA model, especially when the data were not developed according to the assumptions made (e.g., catch-at-age is known without error). It may be beneficial to explore the use of statistical catch-at-age models, which can more easily include the uncertainty in the data and biological processes throughout the model.

5.3.4 Recommendations on estimation modeling

Based on the above findings, there were several recommendations for improving the estimation models used for the JS-T assessment:

1) It is strongly recommended that model diagnostics be included in the assessment documents.

2) It is critical that the uncertainty be treated appropriately in the stock assessment. It is therefore strongly recommended that future assessments do not assume that estimated quantities (e.g., N-at-age, F-at-age, recruitment, spawning biomass) are known without error and, where appropriate, uncertainty

in the data collection and biological processes be included and propagated through the estimation model.

3) It would be beneficial to explore the use of statistical catch-at-age models, which can be more flexible and more easily include the uncertainty in the data and biological processes throughout the model. It would also be easier to develop models with alternative hypotheses.

4) If there are continued difficulties in obtaining data from non-Japanese fisheries, it is recommended to use alternative assessment and management methods for this stock. For example, an observation-based management procedure that sets the catch limits based on a survey of the stock biomass in Japanese waters before the fishing season starts, could be used to manage the stock for Japanese fisheries. However, it would be important to use simulations (e.g., an MSE-type process) to evaluate if such a procedure would meet management objectives and be robust for this stock, including both Japanese and non-Japanese fisheries that are not under the management procedure. Alternatively, data-poor methods can be used to assess the stock but the associated uncertainty would be relatively high and may not be appropriate for the management of this stock.

5.4 Projections

The outputs from the VPA model were used to estimate the stock-recruitment relationship (SRR) and in the projections. Given that the results from the VPA model were not BSIA (Section 5.3), this review also found that the results of the projections were not BSIA. Nevertheless, this review briefly examined the methodology of the projections.

5.4.1 Stock-recruitment relationship (SRR)

It was reported during the web meeting that the assessment scientists follow guidelines from FRA to estimate the SRR in a stock assessment. These guidelines appear to be used across all FRA stock assessments and is beyond the ToRs for this review. Therefore, this review did not consider whether the SRR was BSIA.

Overall, discussion for this stock was similar to the JS-P stock, with a focus on identifying the appropriate regimes for the current stock status and projections. Given the currently available information, this review considered that the use of the ‘normal’ regime to be reasonable. However, it was noted that there was some evidence for a potential transition to a higher recruitment regime. It was recommended that the criteria for classifying the regime be explained clearly in the assessment documents. It was also recommended that the recruitment and SSB estimates used to develop the SRRs have uncertainty associated with them.

5.4.2 Short-term projections

The short-term projections for the JS-T stock were similar to the JS-P stock and had similar problems (see Section 4.4.2 for details). However, unlike the JS-P stock, the terminal year of the estimation model was 2018 (instead of 2019 for the JS-P stock) because of estimation problems for 2019. The allowable catch for the current fishing year was based on SSB in 2021, which necessitated three years of projections to bring forward the estimated N-at-age from 2018. Given that the results from the VPA model were not considered to be BSIA, it follows that the results of these projections were also not considered to be BSIA because the projections rely on the VPA model results for inputs.

In terms of the methodology, the projection models were relatively straightforward and appropriate, given the information available. However, it should be noted that the projected spawning biomass in the current fishing year (i.e., 2021), on which the allowable catch was calculated from, was highly uncertain because the projections were highly sensitive to the uncertain and unobserved estimates of young fish in the population in the terminal year and later (i.e., 2018), changes in recruitment, and catches by non-Japanese fisheries are not limited by Japanese management. It should also be noted that this uncertainty was under-estimated because the N-at-age in the terminal year of the estimation model was assumed to be known without error, which was considered inappropriate (Section 4.3.3). One potential way to improve the short-term projections of the JS-T stock is to develop scientific surveys of this stock in Japanese waters and use the survey results as the basis for managing the stocks in Japanese waters.

5.4.3 Long-term projections

The management objectives and long-term projections for the JS-T stock was similar to the JS-P stock (see Section 4.4.3). Similarly, the results of these long-term projections were not considered to be BSIA because the inputs for the projections relied on the VPA model results, which were not BSIA. The models used for the long-term projections were similar to the short-term projections, and were relatively straightforward and appropriate. However, the uncertainty for the long-term projections were under-estimated because the N-at-age in the terminal year of the estimation model was assumed to be known without error, which was considered inappropriate (see Section 4.3.3). This would be especially important for calculating the probabilities of exceeding the limit and ban reference points, which are based on the tails of the probability distributions. In addition, catches by non-Japanese fisheries are not limited by Japanese management. Therefore, it may not be reasonable to assume that the F-multiplier applies to non-Japanese fisheries as well.

Another important consideration is that of potential changes in recruitment regime over the projection period. Results based on an alternative SRR, with data from the normal recruitment regime but excluding the transition period, were also provided to the managers as additional information. However, unlike the JS-P stock, there does not appear to an alternative recruitment scenario incorporating the recent improved recruitment trend. It was recommended that alternative recruitment scenarios based on the statistics of the previously observed regimes might be useful for managers. There were also results from alternative scenarios that eased the decreases in allowable catch upon the introduction of harvest control rules. However, this review considered these scenarios to be largely management driven and outside the scope of this review.

5.4.4 Recommendations on projections

Based on the above findings, there were several recommendations for improving the SRRs and projections used for the JS-T assessment:

1) It is recommended that the criteria for classifying the regime be described clearly in the assessment documents so that the criteria can be consistently used in future assessments.

2) It is recommended that the recruitment and SSB estimates used to develop the SRRs have uncertainty associated with them.

3) It is recommended to incorporate that uncertainty in the N-at-age in the terminal year into the short term and long-term projections.

4) It is recommended to estimate the statistical distribution and autocorrelation of the estimated recruitment deviates, and use those statistics to represent future recruitment in the long-term projections.

6. Jack Mackerel – Pacific Stock (JM-P Stock)

The stock assessment of the JM-P stock had similarities to the other stocks in this review. Therefore, if similar comments or discussions were made for the other stocks, the reader will be referred to earlier sections to reduce the repetitiveness and lengthiness of the report. However, the recommendations for the JM-P stock were kept separate for the sake of clarity.

6.1 Data

Similar to the other stocks in this review, a VPA model was used to estimate the population dynamics of the JM-P stock during the historical period. A VPA model assumes that the catch-at-age is known without error. Therefore, it was critical to examine the data preparation and determine if this assumption was severely violated. In addition, the relative abundance indices used as tuning indices for the VPA model were also examined.

6.1.1 Catch-at-age

Much of the discussion during the web meeting was focused on the development of the catch-at-age data. Overall, this review found that the catch-at-age data used for the JM-P stock assessment was not considered to be BSIA because the age composition data were likely not BSIA.

In contrast to the other stocks in this review, it was reported that there were negligible amounts of catch by non-Japanese fisheries because the distribution of the JM-P stock is primarily within the Japanese EEZ. However, there were some concerns that the total catch in this assessment likely included some fish from the Jack Mackerel – Tsushima (JM-T) stock (see Section 6.2.1) but the importance of this was currently unknown. Nevertheless, the total catch in weight used in this assessment was considered appropriate but it was strongly recommended to perform more research on the stock structure and distribution, especially the contribution of the JM-T stock to the total catch.

However, several lines of evidence suggested that the age compositions and catch-at-age data were not BSIA. Most importantly, the age composition data of the JM-P stock were based on cohort slicing of length data, which resulted in all fish in a single length bin being assigned to a single age. In addition, a single fixed growth curve was used for all areas and periods, which assumed that growth was constant for all areas and time. Evidence from an unpublished study strongly suggested that this assumption was not valid for this stock, with growth of this stock appearing to vary substantially by area (unpublished data from Watanabe et al.). Secondly, fish from one area likely moved to other areas with different growth rates, which made it very difficult to model growth appropriately. Thirdly, based on the length histograms shown at the web meeting, the modes of the length histograms varied over time and the assumed ages did not consistently match the modes in the length histograms. This indicated that growth likely varied over time and the assumed age distributions were not supported by the length histograms. In addition, the length-weight relationship used to convert catch in weight to catch in numbers was based on a single fixed length-weight relationship for all areas and periods. However, there was no evidence to suggest that this assumption was appropriate.

It is strongly recommended that a robust sampling program for the age composition of the catch using hard parts (e.g., otoliths) be initiated, similar to the JS-P and JS-T stocks. In addition, it is recommended to initiate a study on the spatial and temporal variability of the length-weight relationship of this stock.

6.1.2 Relative abundance indices

Six tuning indices based on data from a variety of fisheries along the Pacific coast were used for the JM-P assessment: 1) CPUE of a set net fishery in southern Miyazaki Prefecture (Index 1), 2) CPUE of a purse seine fishery in Uwajima port (Index 2), 3) CPUE for a medium-scale purse seine fishery in Sukumo Bay (Index 3), 4) catch of age-0 fish in a dip net fishery in Kushimoto (Index 4), 5) catch of age-0 fish from a small-scale trawl fishery in Ise Bay (Index 5), and 6) catch of age-0 fish from a set net fishery in Chiba Prefecture. Based on the discussion during the web meeting, this review did not consider these tuning indices to be BSIA.

All six of the indices were nominal indices (i.e., unstandardized) based on age-0 fish, which assumed that the catchability of all six fisheries have not changed over the entire historical period. However, there was no evidence to support this assumption. It was reported that scientists are currently investigating standardization of these indices. Three of the indices (i.e., Indices 4 – 6) were based on only the catch of the fishery with no effort data, which assumed that the effort of the fisheries were constant over the entire period. However, no evidence was presented to support this assumption. Each of the six fisheries also appeared to be operating over only a small portion of the stock’s range, and therefore did not appear to be representative of the stock’s abundance trends as a whole. Given this, the decision was made to assume that each index was equally representative of the stock, and all six indices were fit in the model with equal weights. Doing so without considering which indices were more representative of the stock and which indices were more consistent with each other and with the catch-at-age data was considered a bad idea because the model will typically just end of going through the middle of all the indices. Overall, the development and use of indices did not appear to be well thought out.

There was also a discussion of the decision to use only age-0 indices. It was reported that age-0 jack mackerel were more valuable than the older ages in the Japanese market and these fish were more heavily targeted and thus provide better indices. Nevertheless, given that the fisheries management was based on the SSB, this review would consider it useful to examine indices for other age classes as well. It is also unclear if targeting results in a better index because it is likely that fishermen adjust fishing operations to improve catchability of a highly targeted stock, such that there is an increased risk of hyperstability in such an index.

It was strongly recommended that future assessments examine the use of survey data like the EPI of the JS-P stock. Given that the fisheries have relatively small operating areas and very different operating characteristics, it might be difficult to find any indices that are representative of the stock as a whole. Based on the experience of mackerel assessments in the EPO, it was beneficial to use the data collected by acoustic-trawl surveys (see Section 4.1.2), even though the surveys were primarily designed for sardines. Therefore, it may be worthwhile to examine if surveys designed for other species have data useful for the JM-P stock as well. If such surveys do not exist and resources are not available for a survey of the stock, it is critical to examine each candidate index carefully and identify the indices most representative of the stock and most amenable to standardization. For example, it should be noted that indices based only on catch without effort data should not be considered unless there is strong evidence that effort has not changed over time, set net indices would also be poor candidates because such indices would be more indicative of changes in movements and distributions of the stock rather than abundance, and purse seine indices tend to be indicative of school size rather than population abundance. Once such standardized indices are identified, the indices should be compared and if need be, different indices can be used to represent scenarios and used in a series of models instead of a single model.

6.1.3 Recommendations on data

Based on the above findings, there are several recommendations for improving the data used for the JM-P assessment:

1) It is strongly recommended to establish a robust sampling and aging program to develop the ALKs and/or age compositions for the assessment. The design of the sampling program is beyond the scope of this review but there would need to be careful consideration of the variability of growth in different areas and periods, as well as selectivities of different fisheries.

2) It is recommended to perform experiments to compare different hard parts for aging and to validate the aging.

3) It would also be recommended to obtain reader-specific estimates of bias and error in the aging.

4) It is recommended to conduct a study to compare the length-weight relationships between months or seasons for different areas to determine if it is important to have finer temporal scale length-weight relationships.

5) It recommended to review the sampling programs in the future.

6) It is strongly recommended that future assessments examine the use of survey data, especially acoustic-trawl data, even if the surveys are designed primarily for other species, to develop indices for this stock.

7) If survey data is unavailable, it is critical that future assessments examine the catchability of these indices and use standardized indices, where appropriate.

8) It is recommended to consider indices for ages-1 to 3 as well as age-0 fish.

9) It is recommended to re-consider the approach taken to develop indices for this stock and how to use the indices in the assessment.

6.2 Biology

6.2.1 Stock structure and distribution

The stock structure for this stock assessment appeared to be appropriate, given the current available information, but more research is needed. The JM-P stock spawns primarily along the Pacific coast of Japan but it was noted that some fish from the JM-T stock moves from the East China Sea into the fishing grounds of the JM-P stock along the Pacific coast and are caught by the JM-P fisheries. Therefore, it was clear that there was some contribution of JM-T stock to the total catch reported in this assessment for the JM-P stock but the amount and importance of that contribution is currently unknown. It was reported that these JM-T fish that migrate to the Pacific coast do not appear to return to the East China Sea to spawn but it is currently unknown if they contribute to the SSB of the JM-P stock. It was strongly recommended that more research be done on the stock structure and distribution of both the JM-P and JM-T stocks, especially on the contribution of the JM-T stock to the catch and potentially SSB of the JM-P stock. It should be noted that the JM-P stock is likely several fold smaller than the JM-T stock and consequently, the JM-T stock has more influence on the JM-P stock than vice versa.

6.2.2 Natural mortality

The M of the JM-P stock was assumed to be 0.5 y^{-1}, using the relationship of $M=2.5/A_{max}$ from Tanaka (1960) and an assumed A_{max} of 5 years. However, it was reported that some aged otoliths from this stock has suggested an A_{max} of 10 years and there has not been a reevaluation of the Tanaka (1960) relationship using more recent data. For similar reasons to the other assessments in this review (see Section 4.2.2), the M for this stock was likely in a reasonable range but it was recommended that the M be re-evaluated and improved. It was therefore strongly recommended that research on M be conducted prior to the next assessment. It would be recommended to perform a meta-analysis (Hamel 2015) using more recent empirical relationships and with more data than Tanaka (1960). In addition, it would be preferable to obtain the maximum age from otolith samples because ages from scale samples tend to biased low, especially for higher ages.

6.2.3 Maturity

The JM-P stock was thought to exhibit 50% maturity at age-1 and be fully mature by age-2. This was considered to be appropriate given the current available information. However, given that the information was from two studies in the 1970s and the variability in growth by area and period, it was recommended to re-examine the maturity of the stock using more recent methods and data, especially focusing on variability between areas and years.

6.2.4 Growth

It was reported that growth of this stock varied substantially by area and period (see Section 6.1.1). Given the variability in growth by area and period, it will be difficult to use growth curves for future assessments. Instead, it was recommended to initiate a sampling program for this stock included sampling of hard parts for aging (see Section 6.1.3).

6.2.5 Recommendations on biology

Based on the above findings, there were several recommendations for improving the biological assumptions and parameters used for the JM-P assessment:

1) It is strongly recommended that more research be done on the stock structure and distribution of both the JM-P and JM-T stocks, especially on the contribution of the JM-T stock to the catch and potentially SSB of the JM-P stock.

2) It is strongly recommended that the natural mortality for this stock be re-examined by performing a meta-analysis using empirical relationships that are more recent and richer in data.

3) It is recommended to estimate the maximum age from otolith samples rather than scale samples because ages from scale samples tend to biased low, especially for higher ages.

4) It is recommended that the maturity be re-examined, especially focusing on variability by area and years.

6.3 Estimation Model

Similar to the other stocks, this review found that the results of the VPA model of the JM-P stock could not be considered as BSIA because the primary data inputs for the assessment were not BSIA. This review did not spend much time on the VPA methodology during the web meeting because it was clear that the data issues identified precluded the VPA results from being considered as BSIA.

6.3.1 VPA model

A standard tuning VPA model was used in the JM-P stock assessment to estimate the historical population dynamics of the stock. This review noted that the strict assumptions about the data for a VPA model have been severely violated in this assessment, which was similar to the other stocks (e.g., Section 4.1). The priority for future improvements would therefore be to solve the problems with the data first. It was also recommended to explore the use of statistical catch-at-age models, especially using an empirical weight-at-age approach (Kuriyama et al. 2015), which can reduce estimation bias for stocks with complex growth compared to modeling parametric growth. Given the complexities of growth for this stock (Sections 6.1.3 and 6.2.4), a carefully designed sampling program for hard parts of the catch and the aging of hard parts is required. Given that non-Japanese fisheries are negligible for this stock, it is likely that data improvements for this stock is plausible in the near future. If research on the stock structure and distribution of this stock indicates that the JM-T stock spawning in the East China Sea has major contributions to the catch of this stock, it would be recommended to examine a joint spatially-explicit model of the JM-T and JM-P stocks, with a portion of the JM-T recruitment moving into the JM-P area.

6.3.2 Model diagnostics

The assessment documents did not provide model diagnostics like retrospective patterns and model fit. However, these model diagnostics for the base case model were shown during the web meeting. The model diagnostics indicated that the indices were poorly fit. There were some retrospective patterns but these patterns were not severe.

6.3.3 Uncertainty

The treatment of uncertainty in the JM-P assessment was similar to the other stocks and was similarly inadequate (see Section 5.3.3 for more details). It was recommended to explore the use of statistical catch-at-age models, which can more easily include the uncertainty in the data and biological processes throughout the model.

6.3.4 Recommendations on estimation modeling

Based on the above findings, there were several recommendations for improving the estimation models used for the JM-P assessment:

1) It is strongly recommended that model diagnostics be included in the assessment documents.

2) It is critical that the uncertainty be treated appropriately in the stock assessment. It is therefore strongly recommended that future assessments do not assume that estimated quantities (e.g., N-at-age, F-at-age, recruitment, spawning biomass) are known without error and, where appropriate, uncertainty in the data collection and biological processes be included and propagated through the estimation model.

3) It would be recommended to explore the use of statistical catch-at-age models, especially with an empirical weight-at-age approach.

4) If research on the stock structure and distribution of this stock indicates that the jack mackerel caught along the Pacific coast has major contributions from the JM-T stock spawning in the East China Sea, it would be recommended to examine a joint spatially-explicit model of the JM-T and JM-P stocks, with a portion of the JM-T recruitment moving into the JM-P area.

6.4 Projections

The outputs from the VPA model were used to estimate the SRR and in the projections. Given that the results from the VPA model were not considered to be BSIA (Section 6.3), this review also found that the results of the projections were not BSIA. Consequently, there was relatively little time spent on reviewing the projections.

6.4.1 Stock-recruitment relationship (SRR)

For similar reasons to the other stocks in this review (see Section 4.4.1), this review did not discuss the SRR substantially during the web meeting and did not consider whether the SRR was BSIA. Overall, given the constraints of the estimation model and data, the approach to estimate the SRR was adequate but left substantial room for improvement. One major issue was that the recruitment and SSB estimates were assumed to be known without error. It was therefore recommended that the recruitment and SSB estimates used to develop the SRRs have uncertainty associated with them. If possible, it would be substantially better to estimate the SRR parameters within the estimation model, rather than using a posthoc approach.

6.4.2 Short-term projections

The terminal year of the estimation model was 2019 but the allowable catch for the current fishing year was based on SSB in 2021. Therefore two years of projections are required to bring forward the estimated N-at-age in 2019 to estimate the projected spawning biomass in 2021. Given that the results from the VPA model were not considered to be BSIA, it follows that the results of these projections were also not BSIA because the projections rely on the VPA model results for inputs.

In terms of the methodology, the projection models were the same as for the other stocks in this review and the comments were similar. One difference was that autocorrelation was included for the projection models. The models were relatively straightforward and appropriate, given the information available. However, it should be noted that the uncertainties in these projections were under-estimated because the N-at-age in the terminal year of the estimation model was assumed to be known without error, which was considered inappropriate. It is therefore important and recommended to incorporate that uncertainty in the N-at-age in the terminal year into the short term projections. The uncertainty in the projected recruitment is also likely under-estimated.

6.4.3 Long-term projections

The management objectives and long-term projections for the JM-P stock were similar to the other stocks in this review. Similarly, the results of these long-term projections were not considered to be BSIA because the inputs for the projections relied on the VPA model results, which were not BSIA. The models used for the long-term projections were similar to the short-term projections, and were relatively straightforward and appropriate. However, the uncertainty for the long-term projections were under-estimated because the N-at-age in the terminal year of the estimation model was assumed to be known without error, which was considered inappropriate.

6.4.4 Recommendations on projections

Based on the above findings, there were several recommendations for improving the SRRs and projections used for the JM-P assessment:

1) It is recommended that the recruitment and SSB estimates used to develop the SRRs have uncertainty associated with them.

2) It is recommended to incorporate that uncertainty in the N-at-age in the terminal year into the short term and long-term projections.

7. Jack Mackerel – Tsushima Stock (JM-T Stock)

The stock assessment of the JM-T stock had many similarities to the other assessments in this review. Therefore, if similar comments or discussions were made for the other stocks, the reader will be referred to earlier sections to reduce the repetitiveness and lengthiness of the report. However, the recommendations for the JM-T stock are kept separate for the sake of clarity.

7.1 Data

Similar to the other stocks in this review (see Section 4.1), a VPA model was used to estimate the population dynamics of the JM-T stock during the historical period. Therefore, it was critical to examine the data preparation, especially for the catch-at-age data.

7.1.1 Catch-at-age

Much of the discussion during the web meeting was focused on the development of the catch-at-age data. Overall, this review found that the catch-at-age data used for the JM-T stock assessment was not BSIA because the total catch in numbers and age composition data were likely not BSIA.

Critically, the time series of total catch used in this assessment did not include large amounts of catch by important non-Japanese fisheries fishing on the same stock. The total catch in this assessment was based on reported catches of the JM-T stock by Japanese and Korean fisheries but did not include catches by vessels from China. Statistics from FAO showed that the JM-T catches by Chinese fisheries were between 200,000 to 400,000 t annually since 2009. In comparison, the sum of Japanese and Korean catches on the JM-T stock were less than 200,000 t annually since 2009. Therefore, this assessment had neglected perhaps the largest source of fishing mortality on the JM-T stock. Doing so likely resulted in biased results, including estimates of SSB, recruitment and fishing mortality (F). Although the catches by Chinese fisheries were uncertain and were not available prior to 2003 and for 2019, there appeared to be enough information to at least develop a range of reasonable historical catch scenarios for the assessment. Therefore, it is critical that future assessments include the catches from all fisheries on this stock.

Based on several lines of evidence, the age compositions for the assessment of the JM-T stock also did not appear to be BSIA. First, the age composition data of the JM-T stock were based on cohort slicing of length data, which resulted in all fish in a single length bin being assigned to a single age. However, aging data for this stock showed that there was substantial overlap of age classes within length bins, especially at larger lengths (Yoda et al. 2014). Although two ALKs were presented during the web meeting, it was clarified that there was no operational aging for this stock and these ALKs were based on fixed growth curves for two areas and only one age was assigned to a single length bin for a single period (i.e., cohort slicing). Second, size observations for an important fishery (large and medium scale purse seine fisheries in Kyushu) appeared to be based on the number of individuals per fish box (i.e., more individuals per box equates to smaller fish) instead of being based on sampling. It was reported during the web meeting that these fish were sorted at the fish market, and were therefore assumed to have negligible variability. However, there were no data to support this assumption. Size data for other fisheries were based on size samples of the catch. Third, at least some age composition data were based on arbitrary size categories like ‘mini’, ‘very small’, ‘small’, ‘medium’, and ‘large’, but there were no data to support these assumed size category-age transition keys. Fourth, an examination of some length histograms of this stock indicated that the assumed age classes did not align consistently with the length histograms. Fifth, the growth of the JM-P stock (unpublished data from Watanabe et al.) was highly variable over space and time, and it was likely the same for JM-T.

Similar to the JM-P stock, it was strongly recommended that a robust sampling program for the age composition of the catch using hard parts (e.g., otoliths) be initiated for this stock. If ALKs are used, it was recommended to have ALKs based on age samples and for each fishery and time period. If it is not possible to obtain length data for some fisheries, it was recommended to develop keys to directly translate size category (e.g., number per box) to age (i.e., age-boxcategory key, ABK) using age samples from these box categories. In addition, it was recommended to initiate a study on the spatial and temporal variability of the length-weight relationship of this stock.

7.1.2 Relative abundance indices

A large number of indices were used for tuning the VPA model: 1) six indices for age-0 fish, 2) three for age-1 fish,3) one for age-2, and 4) one for age-3+. Based on the discussion during the web meeting, this review did not consider these tuning indices to be BSIA.

Overall, the development and use of indices did not appear to be well thought out. Instead, it appeared that the available indices (many of which were conflicting) were just all thrown into the model. For example, in response to a query on the reasons for using so many indices, it was reported during the web meeting that many surveys on this stock had been conducted because the stock structure was complex, and this resulted in many indices being available. However, the approach taken by this assessment for dealing with the reported complexities was inappropriate and likely the opposite of what should be done. Instead of throwing all indices into the model and hoping that the model would sort it out, it was recommended that future assessments come up with a series of hypotheses for the stock's dynamics and assign a small number of consistent indices to represent each hypothesis, resulting in a series of consistent models. The assessment would then be based on an ensemble of consistent models instead of a single model with inconsistent data. In addition, the development of the indices were also inappropriate with some nominal and some standardized indices. It was reported that more standardization will be completed by the next assessment.

It is strongly recommended that a rationalization and review be conducted of the survey designs used for the many surveys for this stock. In general, a single, well-designed survey that covers the entire stock's range each year will be vastly superior to a large number of smaller, uncoordinated surveys. Based on the experience of mackerel assessments in the EPO, it was beneficial to use the data collected by acoustic-trawl surveys (see Section 4.1.2), even though the surveys were primarily designed for sardines. Given that the range of this stock extends to non-Japanese waters, such a survey would be even more critical. If the JM-T catch of non-Japanese fisheries remains poorly known in the future, it may be more appropriate to use alternative assessment and management methods for this stock. For example, an observation-based management procedure that sets the catch limits based on a survey of the stock biomass in Japanese waters before the fishing season starts, could be used to manage the stock for Japanese fisheries (see Section 5.3.1).

7.1.3 Recommendations on data

Based on the above findings, there were several recommendations for improving the data used for the JM-T assessment:

1) It is critical to include catches from all fisheries, including non-Japanese fisheries, that capture this stock of jack mackerel for the assessment.

2) It is strongly recommended to establish a robust sampling and aging program to develop the ALKs and/or age compositions for the assessment. The design of the sampling program is beyond the scope of this review but there would need to be careful consideration of the variability of growth in different areas and periods, as well as selectivities of different fisheries.

3) It is recommended to perform experiments to compare different hard parts for aging and to validate the aging.

4) It would also be recommended to obtain reader-specific estimates of bias and error in the aging.

5) It is recommended to conduct a study to compare the length-weight relationships between months or seasons for different areas to determine if it is important to have finer temporal scale length-weight relationships.

6) It is recommended to review the sampling programs in the future.

7) It is strongly recommended that there be a rationalization and review of the survey design used for the many surveys for this stock. In general, a single, well-designed survey that covers the entire stock's range each year will be vastly superior to a large number of smaller, uncoordinated surveys.

8) It is strongly recommended to examine using an acoustic-trawl survey as the primary source of information for assessing and managing this stock in Japanese waters. Given that the range of this stock extends to non-Japanese waters and the catch of non-Japanese fisheries may be poorly known in the future, an observation-based management procedure that sets the catch limits based on a survey of the stock biomass in Japanese waters before the fishing season starts, could be used to manage the stock for Japanese fisheries.

9) If an acoustic-trawl survey is not applicable, it is strongly recommended that there be a rationalization and review of the many surveys for this stock. In general, a single, well-designed survey that covers the entire stock's range each year will be vastly superior to a large number of smaller, uncoordinated surveys.

7.2 Biology

7.2.1 Stock structure and distribution

The stock structure for this stock assessment appeared to be appropriate, given the current available information. However, it was reported during the web meeting that the structure for this stock is highly complex, which led to numerous indices being developed. In addition, part of the JM-T stock is also caught by JM-P fisheries but the connections between the two stocks are not well known. Therefore, it is strongly recommended that more research be done on the stock structure and distribution of both the JM-P and JM-T stocks.

7.2.2 Natural mortality

For similar reasons to the other assessments in this review (e.g., Section 4.2.2), the M for this study was likely in a reasonable range but it was recommended that the M be re-evaluated and improved. It was recommended to perform a meta-analysis (Hamel 2015) using more recent empirical relationships and with more data than Tanaka (1960). In addition, it would be preferable to obtain the maximum age from otolith samples because ages from scale samples tend to biased low, especially for higher ages.

7.2.3 Maturity

The JM-T stock was thought to exhibit 50% maturity at age-1 and be fully mature by age-2. This was considered to be appropriate given the current available information. However, it was recommended to re-examine the maturity of the stock using more recent methods and data, especially focusing on variability between areas and years.

7.2.4 Growth

It was recommended to initiate a sampling program for this stock included sampling of hard parts for aging, which will also improve the understanding of how growth varies between areas and periods.

7.2.5 Recommendations on biology

Based on the above findings, there were several recommendations for improving the biological assumptions and parameters used for the JM-T assessment:

1) It is strongly recommended that more research be done on the stock structure and distribution of both the JM-P and JM-T stocks.

2) It is strongly recommended that the natural mortality for this stock be re-examined by performing a meta-analysis using empirical relationships that are more recent and richer in data.

3) It is recommended to estimate the maximum age from otolith samples rather than scale samples because ages from scale samples tend to biased low, especially for higher ages.

4) It is recommended that the maturity be re-examined, especially focusing on variability by area and years.

7.3 Estimation Model

Similar to the other stocks, this review found that the results of the VPA model of the JM-T stock were not BSIA because the primary data inputs for the assessment were not BSIA. This review did not spend much time on the VPA methodology during the web meeting because it was clear that the data issues identified precluded the VPA results from being considered as BSIA.

7.3.1 VPA model

A standard tuning VPA model was used in the JM-T stock assessment to estimate the historical population dynamics of the stock. This review noted that the strict assumptions about the data for a VPA model were severely violated in this assessment (see Section 7.1). The priority for future improvements would therefore be to solve the problems with the data first. It was also recommended to explore the use of statistical catch-at-age models, especially using an empirical weight-at-age approach (Kuriyama et al. 2015), which can reduce estimation bias for stocks with complex growth compared to modeling parametric growth.

If research on the stock structure and distribution of this stock indicates that the JM-T stock contributed significantly to the catch of the JM-P stock, it would be recommended to examine a joint spatially-explicit model of the JM-T and JM-P stocks. As suggested in Section 7.1, it was recommended to consider an observation-based management procedure in the future that sets the catch limits based on a survey of the stock biomass in Japanese waters before the fishing season starts.

7.3.2 Model diagnostics

The assessment documents did not provide model diagnostics like retrospective patterns and model fit. However, these model diagnostics for the base case model were shown during the web meeting. The model diagnostics indicated that the indices were very poorly fit and there were moderate retrospective patterns.

7.3.3 Uncertainty

The treatment of uncertainty in the JM-T assessment was similar to the other stocks and was similarly inadequate (see Section 5.3.3 for more details). If a model-based assessment approach is desired, it was recommended to explore the use of statistical catch-at-age models, which can more easily include the uncertainty in the data and biological processes throughout the model.

7.3.4 Recommendations on estimation modeling

Based on the above findings, there are several recommendations for improving the estimation models used for the JM-T assessment:

1) It is strongly recommended that model diagnostics be included in the assessment documents.

2) It is critical that the uncertainty be treated appropriately in the stock assessment. It is therefore strongly recommended that future assessments do not assume that estimated quantities (e.g., N-at-age, F-at-age, recruitment, spawning biomass) are known without error and, where appropriate, uncertainty

in the data collection and biological processes be included and propagated through the estimation model.

3) It would be recommended to explore the use of statistical catch-at-age models, especially with an empirical weight-at-age approach.

4) If research on the stock structure and distribution of this stock indicated that the JM-T stock contributes significantly to the catch of the JM-P stock, it would be recommended to examine a joint spatially-explicit model of the JM-T and JM-P stocks.

7.4 Projections

The outputs from the VPA model were used to estimate the SRR and in the projections. Given that the results from the VPA model were not BSIA (Section 7.3), this review also found the results of the projections to not be BSIA. Consequently, there was relatively little time spent on reviewing the projections.

7.4.1 Stock-recruitment relationship (SRR)

The approach for the SRR of the JM-T stock was similar to the other stocks (e.g., Section 6.4.1), which was based on guidelines from FRA and is beyond the ToRs for this review. Given that, this review did not discuss the SRR substantially during the web meeting and did not consider whether the SRR was BSIA. Overall, given the constraints of the estimation model and data, the approach to estimate the SRR was adequate but left substantial room for improvement. The recommendations would be the same as for the JM-P stock (Section 6.4.1).

7.4.2 Short-term projections

Similar to the JM-P stock, two years of projections were required to bring forward the estimated N-at-age in 2019 (i.e., terminal year of the VPA model) to the projected spawning biomass in 2021, which was used to set the allowable catch. Given that the results from the VPA model were not considered to be BSIA, it follows that the results of these projections were also not considered to be BSIA because the projections relied on the VPA model results for inputs.

In terms of the methodology, the projection models were the same as for the other stocks and the comments were similar (see Section 4.4.2). The models were relatively straightforward and appropriate, given the information available.

7.4.3 Long-term projections

The management objectives and long-term projections for the JM-T stock were similar to the other stocks in this review. Similarly, the results of these long-term projections were not considered to be BSIA because the inputs for the projections relied on the VPA model results, which were not considered to be BSIA. The models used for the long-term projections were similar to the short-term projections, and were relatively straightforward and appropriate. However, the uncertainty for the long-term projections are currently under-estimated because the N-at-age in the terminal year of the estimation model was assumed to be known without error, which was considered inappropriate.

7.4.4 Recommendations on projections

Based on the above findings, there were several recommendations for improving the SRRs and projections used for the JM-T assessment:

1) It is recommended that the recruitment and SSB estimates used to develop the SRRs have uncertainty associated with them.

2) It is recommended to incorporate that uncertainty in the N-at-age in the terminal year into the short term and long-term projections.

8. References

Biseau, A. 1998. Definition of a directed fishing effort in a mixed-species trawl fishery, and its impact on stock assessments. Aquatic Living Resources 11(3):119–136.

Hamel, O. S. 2015. A method for calculating a meta-analytical prior for the natural mortality rate using multiple life history correlates. ICES Journal of Marine Science 72(1):62–69.

Hoenig, J. 1983. Empirical use of longevity data to estimate mortality-rates. Fishery Bulletin 82(1):898–903.

Kuriyama, P., K. Ono, F. Hurtado-Ferro, A. Hicks, I. Taylor, R. Licandeo, K. Johnson, S. Anderson, C. Monnahan, M. Rudd, C. Stawitz, and J. Valero. 2015. An empirical weight-at-age approach reduces estimation bias compared to modeling parametric growth in integrated, statistical stock assessment models when growth is time varying. Fisheries Research 180.

Kuriyama, P. T., J. P. Zwolinski, K. T. Hill, and P. R. Crone. 2020. ASSESSMENT OF THE PACIFIC SARDINE RESOURCE IN 2020 for U.S. MANAGEMENT IN 2020-2021. Page 191.

Lo, N. C. H., B. J. Macewicz, and D. A. Griffith. 2005. Spawning biomass of Pacific sardine (Sardinops sagax), from 1994-2004 off California. Pages 93–112 California Cooperative Oceanic Fisheries Investigations Reports.

Maunder, M. N., and A. E. Punt. 2004. Standardizing catch and effort data: a review of recent approaches. Fisheries Research 70(2–3):141–159.

Okamura, H., Y. Yamashita, and M. Ichinokawa. 2017. Ridge virtual population analysis to reduce the instability of fishing mortalities in the terminal year. ICES Journal of Marine Science 74(9):2427–2436.

Punt, A. E., C. Castillo-Jordán, O. S. Hamel, J. M. Cope, M. N. Maunder, and J. N. Ianelli. 2021. Consequences of error in natural mortality and its estimation in stock assessment models. Fisheries Research 233:105759.

Tanaka, S. 1960. Suisan seibutsu no Population Dynamics to gyogyo shigen kanri (Population Dynamics of fishery organisms and fishery resource management). Bulletin of Tokai Regional Fisheries Research Laboratory 28:1–200.

Then, A. Y., and J. M. Hoenig. 2014. Results of evaluating the performance of empirical estimators of natural mortality rate. Page 26 p. Virginia Institute of Marine Science, College of William and Mary.

Then, A. Y., J. M. Hoenig, N. G. Hall, and D. A. Hewitt. 2015. Evaluating the predictive performance of empirical estimators of natural mortality rate using information on over 200 fish species. ICES Journal of Marine Science 72(1):82–92.

Yoda, M., T. Shiraishi, R. Yukami, and S. Ohshimo. 2014. Age and maturation of jack mackerel Trachurus japonicus in the East China Sea. Fisheries Science 80(1):61–68.

Zwolinski, J. P., D. A. Demer, B. J. Macewicz, G. R. Cutter, B. Elliot, S. A. Mau, D. W. Murfin, R. S Josiah, S. S Thomas, and K. L. Stierhoff. 2016. Acoustic-trawl estimates of northern-stock Pacific sardine biomass during 2015. NOAA Technical Memorandum NOAA-TM-NMFS-SWFSC-559. U.S. Department of Commerce, National Oceanic and Atmospheric Administration, National Marine Fisheries Service, Southwest Fisheries Science Center.

9. Appendix 1 – Terms of Reference

Terms of reference for the Peer Review of Japanese domestic fishery stock assessment

Independent peer review is conducted by external experts on Japanese domestic fishery stock assessment reported by Fishery Research and Education Agency (hereafter, FRA) to evaluate the scientific appropriateness of its assessments. External experts review the appropriateness of the stock assessment and provide recommendation and suggestions for future improvements. FRA either reflects the recommendations to future stock assessment process or provides valid explanation in a document if the suggested proposal is not applicable. The peer review process is performed in accordance with the following tasks:

1. Peer review process is organized by the Secretariat of Peer Review and requires three independent reviewers for each stock: two Japanese reviewers and one overseas reviewer.
2. Japanese reviewers are appointed by the Japanese Society of Fisheries Science (JSFS). Overseas reviewer is appointed from the experts recommended by the National Ocean and Atmospheric Administration (NOAA).
3. In principle, the peer review process is conducted by peer review of the stock assessment report and panel meeting. FRA is required to send stock assessment report to Japanese reviewers few months prior to the panel meeting. For an overseas reviewer, FRA is required to prepare and send English-translated stock assessment report few months prior to the panel meeting. Panel meeting is organized by three independent reviewers, stock assessment team, the Secretariat of Peer Review, and other FRA personnel. In the panel meeting, explanation on stock assessment report is provided by the stock assessment team and question-and-answer session is provided by the independent reviewers.
4. Each independent reviewer prepares a peer review report and submits the document to FRA after peer-reviewing the stock assessment report and attending the panel meeting. A peer review report is roughly 10 pages long and contains the following 7 elements A)-G). After the submission, the Secretariat of Peer Review creates the report of the panel meeting.

    A) Determine whether the data used for stock assessment are adequate to understand the stock dynamics of the target species and represent the best scientific information available (BSIA).

    B) Discuss whether the biological parameters used for stock assessment are appropriate.

C) Discuss whether the basic biological information such as distribution, migration pattern, and population are appropriate.

D) Evaluate whether the stock assessment methodology is based on the most appropriate available study and performed analytically.

E) Evaluate whether the data are treated statistically correctly.

F) Evaluate whether the stock assessment result obtained from the input data and methodology used is appropriate.

G) Evaluate the validity of methodology and result used for the future projection.

5. The Secretariat of Peer Review publishes the peer review report on web page in Japanese and English. Each peer review report is followed by the document of countermeasure that FRA shall take. The proceeding of the panel meeting is also published along with the reports.