ISC/24/SHARKWG-1/1

Mako Age and Growth, Meta-analysis Revisited

Michael Kinney[1] ,Nicholas D. Ducharme-Barth[1], Norio Takahashi[2], Mikihiko Kai[2], Yasuko Semba[2], Minoru Kanaiwa[3], Kwang-Ming Liu[4], José Alberto Rodríguez-Madrigal[5], Javier Tovar-Ávila[5]

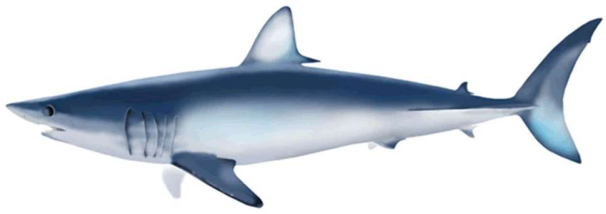
[1] NOAA Fisheries Pacific Islands Fisheries Science Center, 1845 Wasp Boulevard. Honolulu, Hawaii 96818, USA

[2] Fisheries Resources Institute, Japan Fisheries and Education Agency, 5-7-1 Orido, Shimizu, Shizuoka, 424-8633, Japan

[3] Mie University 1577 Kurimamachiya-cho Tsu city, Mie 514-8507 Japan

[4] Institute of Marine Affairs and Resource Management, National Taiwan Ocean University Keelung 20224, Taiwan

[5] Instituto Nacional de Pesca, INAPESCA (National Fisheries Institute of Mexico) Centro Regional de Investigación Pesquera de Bahía Banderas Calle Tortuga #1, La Cruz de Huanacaxtle Nayarit, México 63732



Working document submitted to the ISC Shark Working Group Workshop, 23-25 January 2024, Web-meeting.

Mako Age and Growth, Meta-analysis Revisited

Michael J. Kinney, Nicholas Ducharme-Barth, Norio Takahashi, Mikihiko Kai, Yasuko Semba, Minoru Kanaiwa, Kwang-Ming Liu, José Alberto Rodríguez-Madrigal, Javier Tovar-Ávila

Abstract

Accurate age estimation is a key component for the assessment of any species. Errors in age and growth parameters can have severe consequences both in the accuracy and stability of an assessment model, and in the subsequent management of the species. This is especially true when the species in question is data limited in other ways (catch, effort, etc.) such as with North Pacific shortfin mako shark (mako). Since 2011 members of the ISC Shark Working Group have been seeking to improve our understanding of mako growth by holding growth workshops and collaborating on meta-analysis approaches to best use what data is available. Here we seek to review and better document the information gained and decisions made surrounding the three age and growth workshops held by ISC Shark Working Group members between 2011 and 2017. Additionally we provide an updated age and growth meta-analysis that utilizes laboratory calibration factors for the combination of age data, an updated parametrization of the von Bertalanffy growth model (*Schnute*), and a more statistically appropriate means of combining age readings with length frequency data for an improved ability to estimate key growth parameters. Ultimately our analysis suggests that North Pacific mako are larger at age zero ($L1_{2024}$=65.2; $L1_{2018}$=60) grow slightly slower (k_{2024} female=0.128, male=0.141; k_{2018} female=0.128, male=0.174) and reach a smaller size ($L2_{2024}$ female=272.2, male=225.4; $L2_{2018}$ female=293.1, male=232.8) than what was estimated for the 2018 assessment of North Pacific mako (ISC 2018).

Introduction

The International Scientific Committee for Tuna and Tuna-like Species in the North Pacific Ocean (ISC) completed its first assessment of the North Pacific stock of shortfin mako sharks (here after referred to simply as mako) in 2015 (ISC 2015). The 2015 analysis used a series of fishery indicators, such as catch per unit of effort and average length, to assess the response of the population to fishing pressure. Missing data from important fisheries, and data conflicts (including in age and growth data) led to the conclusion that stock status could not be determined based on the indicator analysis.

Issues with the age and growth of mako were well known in the lead up to the 2015 assessment. For example, basic biological parameters such as asymptotic size ($L\infty$) and length at 50% maturity (L50) had widely varying parameter estimates between studies, even when samples were collected from broadly similar areas of the North Pacific. Estimates for female $L\infty$ ranged from as low as 308.3 cm pre-caudal length (PCL) (Semba, Nakano, and Aoki 2009), to as much as 366.64 cm PCL [converted from original total length using length-length conversion from (Semba, Nakano, and Aoki 2009)] (Hsu 2003). Estimates for female L50 showed similar variations between studies, Joung and Hsu (2005) reported 231.4 cm PCL [converted from original total length using length-length conversion from (Semba, Nakano, and Aoki 2009)], while Semba, Aoki, and Yokawa (2011) reported 256 cm PCL. Knowing this, shark age and growth experts from the ISC Shark Working Group had already held two growth workshops, one in 2011 (ISC 2011)

and one in 2014 (ISC 2014), prior to the completion of the 2015 assessment in an attempt to understand study differences, and promote inter-lab standardization. Another workshop was held in 2017 (ISC 2017b), prior to the next assessment, with the goal of helping to combine data from multiple studies in the North Pacific into a single growth model.

The most recent assessment of mako was carried out in 2018 (ISC 2018), and was a vast improvement over the 2015 analysis. The 2018 assessment was a size-based, age-structured integrated model using the Stock Synthesis modeling platform. One large step forward for the 2018 analysis was the use of a meta-analysis based growth curve (Takahashi et al. 2017). This new analysis was used in order to deal with the identified growth data conflicts from the 2015 assessment, and was a direct product of the preceding three age and growth workshops. Along with improvements to CPUE indices and a more complete accounting of catch, the 2018 assessment was successful in determine a stock status for North Pacific shortfin mako sharks (not overfished, overfishing not occurring).

The purpose of this current working paper is twofold, first, it seeks to review and better document information gained and decisions made surrounding the three age and growth workshops held between 2011 and 2017 that lead to the development of the mako meta-analysis growth model. Second, it looks to add to and improve upon the meta-analysis performed by Takahashi et al. (2017) by incorporating more information from the three age and growth workshops, and also revisiting how the analysis dealt with length frequency data. This work seeks to provide updated age and growth information for North Pacific mako sharks in the lead up to an upcoming assessment in 2024.

Documenting age and growth workshops

First Age and Growth Workshop (2011)

The first workshop of ISC mako age and growth experts was held over two days in December 2011, in La Jolla, California. The main goals of the 2011 workshop were to: 1, discuss the various methodologies used in past regional studies of mako shark age and growth; 2, develop recommendations for standardized methodologies for the collection and processing of shark vertebrae; and 3, begin the development of a vertebra reference collections for cross validation in order to facilitate combining results of past regional studies into a single consensus growth curve. As this was the first meeting of its kind for the working group much of the 2011 workshop focused on the first two goals of understanding and standardizing age and growth methodologies for mako sharks in the North Pacific Ocean.

Some of the advances made during this workshop were the establishment of “best practices” for sample collection/processing, reading, and analysis. The group indicated their preference for vertebra to be collected from just behind the head (vertebrae 15-20), and then frozen prior to cleaning, then stored dry once cleaned, rather than fixed in ethanol. The group agreed to a standard for counting the first band pair (starts at the medial edge of the first narrow/more calcified band), and a suite of measurements to be taken when reading each sample. The group decided that all studies should back calculate from the hypothesized birth month (March: Taiwan; May: Japan, US; July: Mexico) and produce decimal ages to account for growth beyond the end of the year (rather than rounded ages). Finally, the group discussed the pros and cons of each nations processing methods [(*Table 1*) from (ISC 2011)] to help facilitate a future recommendation of which approach to use given the outcome of the reference collection work.

The last major component of the 2011 workshop was to outline the targets for the ISC vertebra reference collection. In reality the group created two collections, the vertebra reference collection being a physical collection of samples that could be split and shared between nations in order to identify laboratory processing differences. The second, a validation collection, was a series of shared digital images form a 2013 validation study (Wells et al. 2013) that the group used to establish consistent counts between labs given the same images which were created using the same laboratory method. While the second, validation collection, was useful in helping the group agree upon counting best practices, like those mentioned above, the information from this was not used to conduct recounts of past studies, or modify ageing techniques for future studies. As such, we mention it here for completeness, but will focus on the vertebra reference collection for the remainder of this work. The group decided on a goal of 50 individual mako (from 70-300 cm) for the reference collection, with an equal split of males and females. As each nations fishery interacted with different sized mako it was decided that responsibilities for supplying vertebra samples would be split depending on the most commonly encountered size range of each fishery: 70-100 cm, USA; 100-150, USA, Mexico; 150-200, Japan; 200-250, Japan; 250-300, Taiwan, Japan (all in total length). Each nation would process their collected samples to the point at which they could be stored dry. Method specific work (e.g. sectioning, staining, etc.) would be done independently by each country once the entire reference collection was shared.

Second Age and Growth Workshop (2014)

The second shark working group age and growth workshop held over three days in January 2014, in La Jolla, California, largely followed the mold of the first, with presentations made on regional age and growth studies, further discussions of methodological differences, and attempts at standardizing research efforts where possible. One major topic that was discussed during the 2014 workshop was that of band pair deposition rates, something which had been a point of discussion since 2013 with the publication of a juvenile mako study (Wells et al. 2013) which indicated that young mako (up to age 5) in the eastern North Pacific appeared to be depositing two growth band pairs per year instead of the one pair per year from previous works in the western North Pacific (Semba, Nakano, and Aoki 2009; Okamura and Semba 2009). While the hypothesis of two band pairs per year did help to partially explain the growth differences seen between eastern and western/central Pacific studies, a general lack of small juvenile mako in eastern Pacific studies made it difficult to resolve band pair deposition rate discrepancies.

Another major component of the 2014 meeting was the distribution of the ISC vertebra reference collection, part of the third goal laid out in the original 2011 workshop. With the original goal of 50 sampled individuals the group was able to collect 65 samples (28 from Japan, 12 from Taiwan, and 25 from the USA) for the vertebra reference collection (34 female, 29 male, and 2 unknown). Each nation received at least two vertebra from every sampled animal so that they could process them using their own standard methodology. Four methods were ultimately used to process reference collection vertebra, centrum-face shadow method (Japan), thin sections with transmitted light (Mexico), soft X-rays (Taiwan), and hard X-rays (USA) (ISC 2014).

Third Age and Growth Workshop (2017)

The third and final age and growth workshop was held over six days in October 2017, in Shimizu, Japan, just prior to the most recent assessment of North Pacific mako. The main topics of the meeting were the shared vertebra reference collection, a discussion of the different ageing techniques used to analyze the

collection and how to interoperate/translate results from each method, as well as a discussion about how best to combine data from relevant past studies into a single meta-analysis.

The group was successful in creating a table indicating how to interoperate the various readings for each method in order to facilitate the translation of one methods counts to another, the interpretation of translucent/opaque bands and edges (in comparison to thick/thin, concave/convex, hypo-mineralized/hyper-mineralized, etc.) [(*Table 2*) from (ISC 2017b)]. The group also discussed how to use and combine age reader's confidence scores. Ultimately, it was decided that no clear method of using this data was known and so confidence scores from different studies were not used in the meta-analysis.

For the vertebra reference collection cross reading project no nation was able to read all 65 samples in the collection (unreadable vertebra based on methods) with the US reading 61, Japan reading 48, and Mexico reading 63 vertebrae. As the only nation with validated age samples (and the largest percent of the vertebra reference collection analyzed) the workshop participants made the decision to treat ages generated by the US method as the basis for which to compare each other method to. As recorded in the workshop report (ISC 2017b), exact agreement between the US counts, and the counts from other methods, was not high (US-Japan 14.9%; US-Mexico 24.6%; US-Taiwan 31.1%) however, the group reported that only the US-Japan counts showed "a significant negative bias (lower counts)" (*Figure 1*).

In terms of the meta-analysis, the above findings lead to a very important decision. The group decided to allow different band pair deposition rates for different ageing methods given "…that cross-reading experiments demonstrated differences in visualization of [band pairs] across studies" (ISC 2017b). Basically, this decision acknowledges a difference in band pair counts when using different laboratory methods, but indicated that no adjustment or correction would be made when combing the age data from these methods in a meta-analysis as the differing band pair deposition rate hypotheses would account for the demonstrated count differences between methods. Therefore, within the meta-analysis, Japan and Taiwan age data would be based on an annual hypothesis, while the US and Mexico age data would be based on a two band pairs down to one following the first five years of growth (i.e. after the first 10 band pairs) hypothesis (Wells et al. 2013; Kinney, Wells, and Kohin 2016).

Another key decision for the meta-analysis following the final growth workshop concerned the use of length frequency datasets. Two length frequency datasets discussed at the workshop (Kai et al. 2015; Runcie et al. 2016) each contained many more records than any of the vertebra based studies (1,000's of records compared to 100's) were included in the analysis. It was decided that due to the scale differences the length frequency datasets would be subset to 125 randomly selected records. Additionally, these records would be confined to animals measuring <150 cm PCL (ages 0-2), as difficulties in distinguishing modes of larger animals created too much uncertainty in age estimation (Takahashi et al. 2017). Importantly, the length frequency datasets where not used as length frequencies in the model but rather as additional sources of age data, as length frequencies were converted to age data using a conversion equation from Kai et al. (2015).

Combining regional studies for population growth estimates

With the conclusion of the third and final growth workshop the task became developing an analysis that could use what was learned about the differences among mako growth studies and combine them into a single model to generate population level growth parameters for a stock assessment.

Growth meta-analysis for North Pacific mako (2017)

Takahashi et al. (2017) set about combining data from seven different sources (five vertebrae based age and growth studies and two length frequency converted age data) into a single von Bertalanffy growth model using a Bayesian hierarchical approach where each growth study was treated as a random effect. With the decision from the age and growth workshops to allow divergent band pair deposition rate hypotheses to account for discrepancies in ageing methodologies between studies, the utility of the ISC vertebra reference collection was simply that of another data source. Subsequently, only the age readings provided by the US (of the ISC vertebra reference collection) were used as one of the five vertebrae based age and growth datasets. This decision also meant that no auxiliary information on the population parameter values of a growth model from previous studies of mako were assumed, meaning that Takahashi et al. (2017) used uninformative priors (diffuse inverse gamma distributions) for both the population parameters and the observation error variances in his model.

In addition to the analysis using the full dataset (7 data sources), two additional sensitives were investigated by Takahashi et al. (2017), one in which all length frequency converted age data were left out of the analysis, and another where age data from a Taiwanese vertebrae ageing study were removed. These sensitives were performed to examine the impacts of these data on resulting parameter estimates as they appeared to have different age and growth characteristics compared to the other growth data in the model. These sensitives indicated that leaving out either the Taiwanese or the length frequency converted age datasets influenced growth model parameter estimates, particularly in terms of estimates of K (leaving out Taiwan data estimated higher K; leaving out length frequency converted age data estimated lower K). The models apparent sensitivity to the selected datasets highlighted the importance of carefully considering model inputs. Ultimately the full dataset, including all seven sources, was deemed the most appropriate for use in the assessment.

While the meta-analysis carried out by Takahashi et al. (2017) was a large step forward in the development of a successful stock assessment (ISC 2018), there still remains room for improvement.

An updated growth meta-analysis for North Pacific mako (2023)

Here we seek to update and add to the original analysis by Takahashi et al. (2017) by altering key aspects of the analysis, such as how the available data were used, how data from different ageing methods were combined in the analysis, how length frequency data were integrated, and the statistical programming language of the model itself.

Materials and Methods

Data used

The vertebral aging data used by Takahashi et al. (2017) are again used here in order to provide consistency between studies. However, unlike in Takahashi et al. (2017) readings from the ISC vertebra reference collection dataset are not used as an additional data set. Rather, readings of the ISC vertebra reference collection from all four labs (US, Japan, Mexico and Taiwan) were used to produce lab specific calibration factors to account for methodological differences between studies.

Certain growth models also included length frequency data via a separate likelihood component (i.e., lengths were not converted to age-at-length data using external growth curves). This allowed growth estimates to be based solely on length modal progression information. In the previous analysis by

Takahashi et al. (2017) subsets of length frequency data from Japan and the US (respectively, (Kai et al. 2015; Runcie et al. 2016)) were used. In the current analysis, given the change in modelling approach for the length frequency data, length measurements from all individuals with identified sex from Runcie et al. (2016) were used. The full length frequency data set from Kai et al. (2015) was not available for the current analysis so the subset of Japanese length frequency data from the Takahashi et al. (2017) was used.

Analysis

Similar to the previous meta-analysis by Takahashi et al. (2017), an integrated Bayesian hierarchical model is used. The previous analysis used Gibbs sampling and the JAGS statistical programming language to sample the posterior distribution. The current analysis applied a more efficient Hamiltonian Monte Carlo (HMC) sampler implemented using the Stan statistical programming language (Stan Development Team 2024b) in R (R Core Team 2024) using the rstan (Stan Development Team 2024a) package. One of the benefits of using Stan is that it provides enhanced diagnostics (e.g., errors and warnings) on how well the HMC sampler is sampling from the posterior distribution. Transitioning from a JAGS model that did not give indication of error to Stan may yield 'new' errors in the Stan model. This is an indication that the JAGS model was inappropriately sampling the posterior distribution, and that modifications to the model are needed in order to appropriately sample the posterior distribution. Analysis progressed in a series of phases, and key sensitivities are noted as sub-headings within each phase:

1. Replicate Takahashi et al. (2017) results using Stan
    a. Estimate using age readings only from 4 labs (exclude ISC reference set and length frequency data)
2. Estimate lab-calibration factors
    a. Assume one of the labs (e.g., US, JP, TW, and MX) has the correct aging methodology
    b. Estimate lab-calibration factors using data only provided by each country in turn
3. Apply lab-calibration factors to standardize the vertebral aging data and estimate von Bertalanffy growth parameters using the Takahashi et al. (2017) model coded in Stan
    a. Assume one of the labs (e.g., US, JP, TW, and MX) has the correct aging methodology
    b. Assume alternative band pair deposition hypothesis (e.g., 1 band pair per year or 2 band pairs per year with a transition to 1 per year at age 5)
4. Reparametrize Takahashi et al. 2017 model to be more consistent with growth assumed in Stock Synthesis (e.g., parametrized in terms of L1, L2, k, CV1, and CV2) and estimate all growth parameters using the standardized vertebral aging data
    a. Exclude TW data
5. Add length frequency likelihood component to estimate growth parameters using both standardized vertebral aging data and length frequency data.
    a. Exclude TW data

Analysis - Phase 1

In the first phase of the analysis the hierarchical von Bertalanffy growth model used by Takahashi et al. (2017) was recoded in Stan. Briefly, the theoretical maximum length ($L_{\infty,j}$) and growth coefficient (k_j) were estimated as random effects for each data set *j*. The model assumed length at age 0 (L_0) was known to be 60 cm. A normal error structure with constant standard deviation at length was estimated for each data-set. All parameters except for L_0 were sex-specific (*s*).

Replicating the model in Stan required a few modifications in order to accommodate the HMC sampler. This included reparametrizing[1] the model to model scaled length (e.g., $l_{i,j}/\max(l_i)$) in order to keep the parameter space on a similar scale, using a non-centered parametrization for the hierarchical random effects, and re-parametrizing the variances to be in terms of standard deviation rather than precision. These are neutral changes resulting in an equivalent model.

The priors used in the Takahashi et al. (2017) analysis were very broad and uninformative. This caused issues with the HMC sampling algorithm (e.g. divergent transition warnings) since the sampler spent too much time in areas of the posterior without any support from the data

Priors – Fixed Effects

$$\mu_{L_\infty,Male} \sim Normal(240, 50)$$

$$\mu_{L_\infty,Female} \sim Normal(310, 50)$$

$$\sigma_{L_\infty,s} \sim Normal^{+}(0, 5)^{2}$$

$$\mu_{k,s} \sim Beta(1, 2.5)$$

$$\sigma_{k,s} \sim Normal^{+}(0, 0.05)$$

$$\sigma_{j,s} \sim Normal^{+}(0, 3)$$

Priors – Random Effects

$$L_{\infty,j,s} \sim Normal\left(\mu_{L_\infty,s}, \sigma_{L_\infty,s}\right)$$

$$k_{j,s} \sim Normal\left(\mu_{k,s}, \sigma_{k,s}\right)$$

Model

$$\tilde{l}_{i,j,s} = L_0 + \left(L_{\infty,j,s} - L_0\right) \times \left(1 - \exp\left(-k_{j,s} \times a_i\right)\right)$$

Likelihood

$$l_{i,j,s} \sim Normal\left(\tilde{l}_{i,j,s}, \sigma_{j,s}\right)$$

Using this model, von Bertalanffy growth parameters were estimated using the five vertebral age and length data sets used in Takahashi et al. (2017). This allowed for a direct comparison with the parameter values reported in Table A3 of Takahashi et al. (2017). An additional sensitivity model was conducted where the ISC reference collection was excluded and only the four vertebral age data sets from the US, Japan, Mexico and Taiwan were used. This allows for direct comparison of the Takahashi et al. (2017) approach with models in the current analysis that are only developed using the four vertebral age data sets from the US, Japan, Mexico and Taiwan.

---

[1] Note that the reparametrizations for rescaling the lengths and non-centered hierarchical random effects are not shown in the Phase 1 equations. The equivalent conventional parametrization of the model is given for ease of interpretability. For example the non-centered parametrization for $L_{\infty,j,s}$ would be: $\ddot{L}_{\infty,j,s} \sim Normal(0,1)$; $L_{\infty,j,s} = \sigma_{L_\infty,s} \times \ddot{L}_{\infty,Male} + \mu_{L_\infty,s}$.

[2] Note that $Normal^{+}$ describes a Normal distribution bounded on the interval $[0, \infty)$

Analysis – Phase 2

The second phase of the analysis consisted of estimating lab-calibration factors using the paired readings in the ISC reference collection from the US, Japan, Mexico, and Taiwan. This data set is previously described above in the section Second Age and Growth Workshop (2014). The model assumed a linear relationship between the latent 'true age' a_i and the observed age from each lab $\tilde{a}_{i,j}$. All labs were assumed to be able to identify age 0 individuals without error. Separate lab-calibration factors β_j and observation error σ_j were estimated for each lab. A Normal prior for the 'true age' a_i was used where the prior mean μ_{a_i} for each a_i was taken as the mean band pair reading between labs ($\mu_{a_i} = \frac{1}{4}\sum_{j=1}^{4} \tilde{a}_{i,j}$), and the prior standard deviation σ_{a_i} was calculated from the prior mean assuming a CV of 0.2 ($\sigma_{a_i} = 0.2 \times \mu_{a_i}$).

Priors – Fixed Effects

$$\beta_j \sim Normal^{+}(1, 0.2)$$

$$\sigma_j \sim Lognormal(0, 0.2)$$

Prior – Random Effect

$$a_i \sim Normal^{+}(\mu_{a_i}, \sigma_{a_i})$$

Likelihood

$$\tilde{a}_{i,j} \sim Normal(\beta_j \times a_i, \sigma_j)$$

The inverse of the lab-calibration factor ($1/\beta_j$) could then be multiplied to the band pair counts from each of the 4 vertebral age-length data sets to develop standardized band pair counts for each of the age-length data-sets.

A set of sensitivity analyses were conducted where the lab-calibration factor β_j for each lab was held fixed at 1. This allowed us to derive calibration factors for the other labs conditioned on the assumption that the lab where $\beta_j = 1$ was the 'true' unbiased aging method.

A second set of sensitivity analyses were conducted where lab-calibration factors β_j were derived in turn from samples contributed by the US, Japan, and Taiwan. This sensitivity was conducted to explore the hypothesis that methodological differences between the labs can be used to explain the differences in band pair counts between labs, and the two competing band pair deposition hypotheses. It was also proposed that the competing band pair hypotheses are valid and a product of spatial variation in biological processes resulting in a single band pair per year being deposited in the Western Pacific Ocean and two band pairs per year being deposited in the Eastern Pacific Ocean up to age 5. If the β_j differ within lab across the different data subsets then this would support the biological variation hypothesis. However, if the β_j are consistently estimated within lab across the different data subsets then this would support the hypothesis that the two different band pair deposition hypotheses are an artefact of methodological differences.

Analysis – Phase 3

Analysis – Phase 3 used the same growth model described in *Analysis - Phase 1* in order to develop growth models under alternative assumptions of band pair deposition and 'true' aging methodology. For this phase in the analysis, the 'ages' used in the Takahashi et al. (2017) study were converted to standardized

band pair counts. Only the data-sets from Japan, Taiwan, the US, and Mexico were used. For the Japanese and Taiwanese data which assumed a one band pair per year hypothesis, the 'ages' were rounded to the greatest integer less than or equal to the initial 'age' (e.g. the *floor* function) before being multiplied by the inverse of the lab-calibration ($1/\beta$) to yield standardized band pair counts. US and Mexico 'ages' assumed a transition from two band pairs per year to one band pair per year at age 5. For 'ages' greater than or equal to 5, 5 was added to the 'age' before applying the *floor* function and multiplying by $1/\beta$ to yield standardized band pair counts. For 'ages' less than 5, the 'age' was multiplied by 2 before applying the *floor* function and multiplying by $1/\beta$. Following, standardization the band pair counts were converted back into age (e.g., standardized age) according to one of the two band pair deposition hypotheses.

Two scenarios were considered for developing the standardized ages used to estimate growth. The first scenario applied the Japanese lab-calibration factor β_{Japan} and assumed a one band pair per year hypothesis. The second scenario applied the US lab-calibration factor β_{US} and assumed a two band pair per year until age 5 hypothesis. These two scenarios represent the two most plausible alternative hypotheses, and bound the differences in ageing across methodologies. The two band pair per year until age 5 hypothesis has been validated using oxytetracycline marked individuals (Wells et al. 2013; Kinney, Wells, and Kohin 2016) and is consistent with the US ageing approach which detected the greatest number of band pairs on average. The one band pair per year hypothesis is consistent with the Japanese aging approach which detected the fewest band pairs on average.

Analysis – Phase 4

The fourth phase of the analysis reparametrized the von Bertalanffy growth model used in Phases 1 & 3 to be more consistent with the formulation of the von Bertalanffy growth model used within Stock Synthesis. The *Schnute* parametrization of growth, using L_1 and L_2, was applied to define the growth curve. These are the lengths associated with ages A_1 and A_2, where $A_1 = 0$ and $A_2 = 40$. The A_1& A_2 values are arbitrarily set to values which ease the interpretability of the estimated parameters and have no influence on the estimated shape of the growth curve. Another benefit of the *Schnute* parametrization is that it is more numerically stable than alternative parametrizations of von Bertalanffy growth. In the Phase 4 model, sex-specific L_2 was estimated as a random effect for each lab *j*, while a single L_1 value was estimated and shared across labs *j* and sexes *s*. Furthermore, rather than parametrize the variance around the growth curve in terms of standard deviations, it is given in terms of CV_1 & CV_2 which are the CVs associated with L_1 & L_2. This is a change to make the estimated growth parameters internally consistent with the parametrization applied within Stock Synthesis. These CVs are converted to standard deviation for each observation according to the following equation:

$\sigma_{i,j,s} = \tilde{l}_{i,j,s}\left(CV_{1,j,s} + \frac{(\tilde{l}_{i,j,s}-L_1)}{(L_{2,j,s}-L_1)}\left(CV_{2,j,s} - CV_{1,j,s}\right)\right)$. The full model is given below noting that as in Phase 1 the reparametrizations for rescaling the lengths and non-centered hierarchical random effects are not shown for ease of interpretability.

Priors – Fixed Effects

$$L_1 \sim Normal(60, 0.6)$$

$$\mu_{L_2,Male} \sim Normal(240, 50)$$

$$\mu_{L_2,Female} \sim Normal(310, 50)$$

$$\sigma_{L_2,s} \sim Normal^{+}(0, 30)$$

$$\mu_{k,s} \sim Normal(0.12, 0.02)$$

$$\sigma_{k,s} \sim Normal^{+}(0, 0.05)$$

$$CV_{1,j,s} \sim Lognormal(\log(0.1), 0.3)$$

$$CV_{2,j,s} \sim Lognormal(\log(0.1), 0.3)$$

Priors – Random Effects

$$L_{2,j,s} \sim Normal(\mu_{L_2,s}, \sigma_{L_2,s})$$

$$k_{j,s} \sim Normal(\mu_{k,s}, \sigma_{k,s})$$

Model

$$\tilde{l}_{i,j,s} = L_1 + (L_{2,j,s} - L_1)\frac{(1 - e^{-k_{j,s}(a_i - A_1)})}{(1 - e^{-k_{j,s}(A_2 - A_1)})}$$

Likelihood

$$l_{i,j,s} \sim Normal(\tilde{l}_{i,j,s}, \sigma_{i,j,s})$$

As in Analysis – Phase 3, the 4 age-length data sets from Japan, Taiwan, the US and Mexico were analyzed. Additionally the same two scenarios for converting the ages to standardized ages were used: β_{US} and the two-band pair per year hypothesis; and β_{Japan} and the one-band pair per year hypothesis. Lastly, a sensitivity analysis was conducted where the Taiwanese age-length data were excluded from the analysis.

Analysis – Phase 5

The last phase of analysis takes the von Bertalanffy growth model described in Phase 4 and adds a dedicated likelihood component for length frequency data inspired by Zhou et al. (2019). In this model, modes in the length frequency data $\breve{l}_i$ are modeled using a mixture model, where the observed modes of the mixture distribution $L_{obs_{t,j,s}}$ are constrained according to the underlying von Bertalanffy growth curve. An additional parameter is needed which corresponds to the age a_0 at the first observed mode $L_{obs_{1,j,s}}$. The number of modes in the mixture distribution T is determined *a priori* based on an inspection of the length frequency distribution. Two layers of random effects are used to model both the process error $\sigma_{VB_{t,j,s}}$ and the observation error $\sigma_{mix_{t,j,s}}$ in the deviation of the observed modes from the underlying von Bertalanffy growth model $L_{VB_{t,j,s}}$. The following equations, in addition to all of those from Analysis – Phase 4, are used to define the full model used in this phase of the analysis.

Priors – Fixed Effects

$$\theta_{raw_t} \sim Normal(0, 1)$$

$$a_0 \sim Normal^{+}(0, 0.25)$$

$$CV_{mix_{j,s}} \sim Lognormal(\log(0.1), 0.3)$$

Priors – Random Effects

$$L_{true_{t,j,s}} \sim Normal\left(L_{VB_{t,j,s}}, \sigma_{VB_{t,j,s}}\right)$$

$$L_{obs_{t,j,s}} \sim Normal\left(L_{true_{t,j,s}}, \sigma_{mix_{t,j,s}}\right)$$

Model

$$\theta_t = \frac{e^{\theta_{raw_t}}}{\sum_{t=1}^{T} e^{\theta_{raw_t}}} : t \in T$$

$$a_t = a_0 + t - 1 : t \in T$$

$$L_{VB_{t,j,s}} = L_1 + \left(L_{2,j,s} - L_1\right)\frac{\left(1 - e^{-k_{j,s}(a_t - A_1)}\right)}{\left(1 - e^{-k_{j,s}(A_2 - A_1)}\right)} : t \in T$$

$$\sigma_{VB_{t,j,s}} = L_{VB_{t,j,s}}\left(CV_{1,j,s} + \frac{\left(L_{VB_{t,j,s}} - L_1\right)}{\left(L_{2,j,s} - L_1\right)}\left(CV_{2,j,s} - CV_{1,j,s}\right)\right)$$

$$\sigma_{mix_{t,j,s}} = L_{true_{t,j,s}} \times CV_{mix_{j,s}}$$

Likelihood

$$\breve{l}_{i,j,s} \sim \sum_{t=1}^{T} \theta_t \times Normal(L_{obs_{t,j,s}}, \sigma_{mix_{t,j,s}})$$

In this model, the length-frequency data $\breve{l}_{i,j,s}$ are included as additional data sets *j* such that a new set of random effects for L_2, k, CV_1, and CV_2 are estimated for each additional length-frequency data set. Two length-frequency data sets were evaluated for inclusion in the model: all individuals with identified sex from (Kai et al. 2015; Runcie et al. 2016) and the subset of Japanese length frequency data from the Takahashi et al. (2017) was used.

As in Analysis – Phase 3 & 4, the 4 vertebral age-length data sets from Japan, Taiwan, the US and Mexico were analyzed. Additionally the same two scenarios for converting the ages to standardized ages were used: β_{US} and the two-band pair per year until age 5 hypothesis; and β_{Japan} and the one-band pair per year hypothesis. Lastly, a sensitivity analysis was conducted where the Taiwanese age-length data were excluded from the analysis.

Stan Diagnostics

As mentioned in a previous section, Stan provides a number of diagnostics to evaluate how well the HMC sampler is mapping the posterior distribution. A divergence is an indication that the step size being used by the sampler is too large for the underlying distribution, and that inference on the resulting posterior distribution could be unreliable. An iteration where the maximum tree depth is reached is an indication that the HMC algorithm is inefficient due to the step size chosen being too small for the distribution. Chains with a low level of energy Bayesian fraction of missing information (E-BFMI) indicates that the HMC sampler had difficulty appropriately mapping the posterior distribution during the 'warm-up' (i.e. burn-in) period.

Stan also reports the value of $\hat{R}$ which can be used to identify if sampling chains are well-mixed and have converged to a stable distribution ($\hat{R} < 1.05$).

Calculating the expected log pointwise predictive density (ELPD) using cross-validation (e.g., *elpd_loo*) can allow us to evaluate if the model is over-parametrized. If the effective number of parameters in the model (*p_loo*) is greater than the number of observations in the model or the number of leading parameters then this may indicate that the model is over-parametrized or mis-specified. Lastly, the Pareto *k* diagnostic is calculated for each observation and is a measure of how far any of the leave-one-out predictive samples are from the full distribution. If *k > 0.7* then this is an indication that the model has a difficult time fitting to that observation and could be an indication that the model is mis-specified or that the observation is an outlier/data-entry error.

Results

Results – Phase 1

The Takahashi et al. 2017 von Bertalanffy growth model was implemented in Stan using 6 chains each with 320 'warm-up' (i.e. burn-in) iterations and 1,280 sampling iterations. Thinning of the posterior chains was not conducted. The HMC sampler assumed a target average proposal acceptance probability of 0.99 (adapt_delta = 0.99) during the warm-up period and a maximum tree depth of 12.

Sampling was completed with no divergences, no iterations saturating the maximum tree depth, and no chains exhibiting pathological behavior according to E-BFMI. The maximum $\hat{R}$ value across all leading parameters was 1.0038 indicating that the chains are well-mixed and converged to a stable distribution.

The *p_loo* value of 30.3 was less than the number of leading parameters 38 indicating that the model is unlikely to be over-parametrized. There was 1 data-point out of 1,039 input data-points with Pareto-*k* > 0.7 indicating that the model provides a good representation of the observations.

As indicated in Figure 2 the growth curves estimated within Phase 1 using all 5 data-sets (Japan, Taiwan, the US, Mexico and the ISC-reference collection) are almost identical to the growth curves based on median values listed in Table A3 of Takahashi et al 2017. This indicates that the Stan version of the model is able to replicate the previous analysis when fitting to the same data. Figure 3 shows the comparison between the Takahashi et al. 2017 growth curve fit to all 7 data sets (blue) which was used in the last assessment relative to a growth curve excluding the length frequency data sets (green), and excluding both the length frequency data sets and the ISC reference data set (orange). Since most ensuing models only fit to the 4 main age-length data sets the most appropriate comparison to the Takahashi et al. 2017 analysis should be made using the orange curve.

Results – Phase 2

The lab calibration model was implemented in Stan using 6 chains each with 320 'warm-up' (i.e. burn-in) iterations and 1,280 sampling iterations. Thinning of the posterior chains was not conducted. The HMC sampler assumed a target average proposal acceptance probability of 0.8 during the warm-up period and a maximum tree depth of 10.

Sampling was completed with no divergences, no iterations saturating the maximum tree depth, and no chains exhibiting pathological behavior according to E-BFMI. The maximum $\hat{R}$ value across all leading parameters was 1.0042 indicating that the chains are well-mixed and converged to a stable distribution.

The *p_loo* value of 67.8 was less than the number of leading parameters 70 indicating that the model is unlikely to be over-parametrized. There were 24 data-points out of 310 input data-points with Pareto-*k* > 0.7 indicating that the model provides a reasonable representation of most of the observations.

Estimated lab-calibration factors β_j and observation error σ_j are shown in Figure 4 when no lab is assumed to have the true methodology and when each lab is in turn assumed to have the true methodology. Relative β_j across labs are consistently estimated across scenarios where the US methodology shows the highest counts and the Japanese methodology shows the lowest counts. When the reference data were restricted to different subsets of data (e.g., Japan samples only, Taiwan samples only, Japan & Taiwan samples, and US samples only) a consistent pattern is seen between the estimation of β_j for Japan and the US (Figure 5). This indicates that it is perhaps methodological differences rather than regional differences in the biology that result in the differing counts between labs, and that the alternative band pair hypotheses may be an artefact of these methodological differences. For the Taiwan data subset the Japanese estimate is centered on the β_j prior since there were no Japanese readings for this subset of data.

Results – Phase 3

The Stan version of the Takahashi et al. 2017 model (described in Phase 1) using data standardized using β_{US} and the two-band pair per year hypothesis was implemented in Stan using 3 chains each with 320 'warm-up' (i.e. burn-in) iterations and 480 sampling iterations. Thinning of the posterior chains was not conducted. The HMC sampler assumed a target average proposal acceptance probability of 0.99 during the warm-up period and a maximum tree depth of 12. Sampling was completed with no divergences, no iterations saturating the maximum tree depth, and no chains exhibiting pathological behavior according to E-BFMI. The maximum $\widehat{R}$ value across all leading parameters was 1.0073 indicating that the chains are well-mixed and converged to a stable distribution. The *p_loo* value of 29.2 was less than the number of leading parameters 32 indicating that the model is unlikely to be over-parametrized. There were 0 data-points out of 982 input data-points with Pareto-*k* > 0.7 indicating that the model provides a good representation of the observations.

The Stan version of the Takahashi et al. 2017 model (described in Phase 1) using data standardized using β_{Japan} and the one-band pair per year hypothesis was implemented in Stan using 3 chains each with 320 'warm-up' (i.e. burn-in) iterations and 480 sampling iterations. Thinning of the posterior chains was not conducted. The HMC sampler assumed a target average proposal acceptance probability of 0.99 during the warm-up period and a maximum tree depth of 12. Sampling was completed with no divergences, no iterations saturating the maximum tree depth, and no chains exhibiting pathological behavior according to E-BFMI. The maximum $\widehat{R}$ value across all leading parameters was 1.0113 indicating that the chains are well-mixed and converged to a stable distribution. The *p_loo* value of 24.7 was less than the number of leading parameters 32 indicating that the model is unlikely to be over-parametrized. There were 2 data-points out of 982 input data-points with Pareto-*k* > 0.7 indicating that the model provides a good representation of the observations.

Growth curves estimated using data standardized using β_{US} and the two-band pair per year hypothesis resulted in a lower L_∞ and higher k than growth curves estimated using data standardized using β_{Japan} and the one-band pair per year hypothesis (Figure 6). Growth curves estimated using data standardized using β_{Japan} and the one-band pair per year hypothesis were most similar to the curve assumed in the

previous stock assessment. The effect of the standardization is seen when comparing the gray dotted line in the 'Aggregate' panel as this curve is fit to the unstandardized data using the same model.

Results – Phase 4

The Phase 4 model using data standardized using β_{US} and the two-band pair per year hypothesis was implemented in Stan using 3 chains each with 320 'warm-up' (i.e. burn-in) iterations and 480 sampling iterations. Thinning of the posterior chains was not conducted. The HMC sampler assumed a target average proposal acceptance probability of 0.99 during the warm-up period and a maximum tree depth of 12. Sampling was completed with no divergences, no iterations saturating the maximum tree depth, and no chains exhibiting pathological behavior according to E-BFMI. The maximum $\widehat{R}$ value across all leading parameters was 1.0095 indicating that the chains are well-mixed and converged to a stable distribution. The *p_loo* value of 29.9 was less than the number of leading parameters 41 indicating that the model is unlikely to be over-parametrized. There were 2 data-points out of 982 input data-points with Pareto-*k* > 0.7 indicating that the model provides a good representation of the observations.

The Phase 4 model using data standardized using β_{Japan} and the one-band pair per year hypothesis was implemented in Stan using 3 chains each with 320 'warm-up' (i.e. burn-in) iterations and 480 sampling iterations. Thinning of the posterior chains was not conducted. The HMC sampler assumed a target average proposal acceptance probability of 0.99 during the warm-up period and a maximum tree depth of 12. Sampling was completed with no divergences, no iterations saturating the maximum tree depth, and no chains exhibiting pathological behavior according to E-BFMI. The maximum $\widehat{R}$ value across all leading parameters was 1.0068 indicating that the chains are well-mixed and converged to a stable distribution. The *p_loo* value of 31.5 was less than the number of leading parameters 41 indicating that the model is unlikely to be over-parametrized. There was 1 data-point out of 982 input data-points with Pareto-*k* > 0.7 indicating that the model provides a good representation of the observations.

Increasing the flexibility of the growth model by estimating L_1, and CV_2 had a noticeable effect on the estimated growth curves (Tables 3 & 4). Looking at the likelihood fits by data component, the β_{US} and the two-band pair per year hypothesis model was able to achieve slightly better fits to all components except the Taiwanese data (Table 5). Relative to the Phase 3 model, estimates of length at age are closer between the two scenarios for the largest sizes (Figure 7). Relative to the Phase 3 models, the Phase 4 models are able to provide a better fit to the data according to elpd_loo (Table 6) with the β_{Japan} and one-band pair per year hypothesis model providing the best overall fit. However, as seen with an examination of the likelihoods by data component this seems to be driven by the fit to the Taiwanese data since removing this component shows that the β_{US} and the two-band pair per year hypothesis model has the best overall fit (Table 7). Excluding the standardized Taiwanese data had a small impact on the estimated growth curves under both scenarios, most notably manifested as a slight increase in the maximum size at length (Figure 8).

Results – Phase 5

A number of models were attempted in Phase 5 to find a configuration that yielded satisfactory convergence diagnostics. Models were fit using Takahashi et al. (2017) subsets of length frequency data from Japan and the US (Figure 9) however these models had poor HMC posterior sampling diagnostics. The full length composition data from all individuals with identified sex from Runcie et al. (2016) (Figure 10) was used on its own and in combination with the subset of Japanese length frequency data available at the time of the analysis. Multiple assumptions were made for the number of modes in the mixture

model (e.g., 3, 4 or 5), the band pair hypothesis, and the inclusion of the Taiwanese data. Ultimately, the only model that showed satisfactory diagnostics was the model assuming 4 modes, included the full Runcie et al. 2016 length composition without the Japanese length frequency data, used the vertebral age data standardized using β_{US} and the two-band pair per year hypothesis, and included the Taiwanese vertebral aging data. The results from this model are presented in further detail.

The Phase 5 model was implemented in Stan using 3 chains each with 320 'warm-up' (i.e. burn-in) iterations and 480 sampling iterations. Thinning of the posterior chains was not conducted. The HMC sampler assumed a target average proposal acceptance probability of 0.8 during the warm-up period and a maximum tree depth of 10. Sampling was completed with no divergences, no iterations saturating the maximum tree depth, and no chains exhibiting pathological behavior according to E-BFMI. The maximum $\widehat{R}$ value across all leading parameters was 1.0090 indicating that the chains are well-mixed and converged to a stable distribution. The *p_loo* value of 57.1 was less than the number of leading parameters 77 indicating that the model is unlikely to be over-parametrized. There was 1 data-point out of 2,223 input data-points with Pareto-*k* > 0.7 indicating that the model provides a good representation of the observations.

Incorporating the full Runcie et al. 2016 length frequency data-set had a minimal impact to the estimated growth curve (Figure 11), as estimated values are very similar to the Phase 4 using data standardized using β_{US} and the two-band pair per year hypothesis (Table 8). Fits to the individual vertebral age data components appear reasonable, though there does appear to be a small number of large individuals for Japan and Taiwan that the model is unable to capture (Figure 11).

The model also appears to do a good job predicting the distribution of length frequency data (Figure 12). The mixture model estimated the age associated with the first mode to be between 0.28-0.35 (Table 9). Given that the majority of sampling typically occurred in July this corresponds to an average pupping date of mid-to-late March, which agrees with estimates from the literature of parturition occurring sometime between late-winter to mid-spring in both hemispheres of the Pacific. For both males and females, the majority of individuals were estimated to be ages 1+ & 2+ (modes 2 & 3; $\theta_{Male,2} = 0.509, \theta_{Male,3} = 0.349, \theta_{Female,2} = 0.573$ & $\theta_{Female,3} = 0.261$). This corresponded to observed modes L_{obs} of 91 – 118 cm PCL for males and 91 – 119 cm PCL for females.

Overall, the model produces similar fits to the β_{US} and the two-band pair per year hypothesis for Phases 3 & 4, and produces growth curves with smaller maximum size and slightly slower growth relative to the curve from Takahashi et al. 2017.

Discussion

Estimating growth for North Pacific mako is a challenging endeavor, their expansive range, scant targeted fishing effort, and the general absence of large adults in the catch of any fishery, means that samples are often spatially restricted (in comparison to the range of the stock), and contain only a limited range of lengths and ages. Collaborative scientific efforts involving several nations' fisheries is required to better understand the fundamental life history of this species. Since 2011 members of the ISC shark working group have been seeking to improve our understanding of mako growth by promoting standardized practices for sample collection/processing (ISC 2011), consistent terminology (ISC 2014), and reader agreement between laboratories (ISC 2017a). Leveraging off of this work Takahashi et al. (2017) represents the first real attempt at a meta-analysis that combined length and age data for mako from

multiple nations in order to try and overcome some of the aforementioned limitations. By building upon Takahashi et al. (2017) original work we have successfully been able to produce an updated meta-analysis for North Pacific mako that utilizes laboratory calibration factors for the combination of age data, an updated parametrization of the von Bertalanffy growth model (*Schnute*), and a more statistically appropriate means of combining age readings with length frequency data for an improved ability to estimate key growth parameters.

The identification of a consistent pattern of dissimilar age estimates between labs using the 2014 vertebra reference collection is a key finding of this new analysis. Rather than solely relying on the possible spatial variability of band pair deposition rates between eastern and western North Pacific mako we are now able to more directly account for laboratory differences prior to combining age data into a meta-analysis. This is a much improved use of the vertebra reference collection and its unique dataset of multiple readings from vertebra of the same individuals processed using different laboratories methods. However, differences in k and L2 by laboratory even after aging estimates were standardized assuming a 'true' methodology and consistent band pair deposition hypothesis suggest that there is still variation unaccounted for between those groups of samples. This could be additional differences in methodology that are not completely captured by the lab calibration factors used to standardize the ages (perhaps due to the limited size of the vertebra reference collection), or it could be indicative of spatial differences in growth between regions where the samples were collected. Nevertheless, the use of the vertebra reference collection to create laboratory calibration factors is a significant improvement over simply using the reference collection as an additional vertebra age dataset.

The use of raw length frequency data (not subset or converted into ages) is also a significant improvement of this updated meta-analysis. By not converting length frequency data to ages we are avoiding any reliance on an external growth estimate to create a dataset. Additionally, by not subsetting the length frequency data we are allowing more information to be passed to the model then if only a few hundred samples from only young individuals are used. However, the aggregation of length frequency data into a single distribution per data set assumes that modal structure is consistent across sampling events and that sampling occurred at the same time relative to pupping in each year. A more appropriate way of dealing with the length frequency data would be to split up the aggregate distribution into data from individual sampling events. This would be more consistent with some of the assumptions made and may result in better fits to the data. Additionally, it would also make the length frequency data more informative for the estimation of the growth parameters. However, this would come at the cost of greatly increased model complexity (e.g., a separate set of mixture model probabilities, age at first mode, and modal random effects would be needed for each sample). This is an area of further development for this model as it is currently not possible to define a different number of mixture model modes for each sampling event.

By updating Takahashi et al. (2017) to include laboratory calibration factors, a reparametrized von Bertalanffy growth model, and a more statistically appropriate means of incorporating length frequency data, this update provides a positive step forward both for the estimation of key growth parameters, and their associated uncertainty. Based on the Phase 5 model that showed satisfactory diagnostics, this analysis suggests that North Pacific mako are larger at age zero ($L1_{2024}$=65.2; $L1_{2018}$=60) grow slightly slower (k_{2024} female=0.128, male=0.141; k_{2018} female=0.128, male=0.174) and reach a smaller size ($L2_{2024}$ female=272.2, male=225.4; $L2_{2018}$ female=293.1, male=232.8) than what was estimated for the 2018 assessment (ISC 2018). While our growth parameter estimates differ from the 2018 assessment they are

within the range of expected values for mako in the North Pacific based on the scientific literature (Semba, Nakano, and Aoki 2009; Hsu 2003; Cailliet and Bedford 1983; Ribot-Carballal, Galván-Magaña, and Quiñónez-Velázquez 2005; Wells et al. 2013). We suggest that the growth estimates and their associated variances (Table 8) from this updated meta-analysis be utilized in the next assessment of North Pacific mako. Additionally, given the variability seen between growth curves associated with each laboratory, we suggest that lab specific growth curves from the phase 5 analysis be considered as alternative growth curves either as sensitivity runs or alternative growth hypotheses within a model ensemble.

Acknowledgements

Thank you to Rob Ahrens (PIFSC) for helpful discussions regarding the set-up of the lab-calibration model, and other components of the revised growth analysis. Also thank you to the members of the ISC Shark Working Group for sharing there mako growth data and their knowledge of the various biology workshops held since 2011.

Tables

Table 1: Results of 2011 workshop group discussion of the pros and cons of all ageing methods currently being used in the study of mako sharks in the north Pacific.

Technique	Pros	Cons	Equipment Needed	Notes
Thin (microtome) sectioning	Easy		Microtome, Microscope	
Staining process added to thin sectioning	Improves upon thin sectioning	More labor intensive than simple thin sectioning	Microtome, Microscope	
Whole centrum with silver nitrate		Chemical disposal issues	Microscope	See studies by Semba and Ribot-Carballal
Hard X-ray (sectioning)	Relatively easy, clear images	Chemical disposal issues	Microtome, X-ray and processor	See studies by Wells, Hsu, Acuña and Cailliet; consider performance of hard vs. soft X-ray
Soft X-ray (whole centra)	Relatively easy, less clear images	Chemical disposal issues	Microtome, X-ray and processor	See studies by Wells, Hsu, Acuña and Cailliet; consider performance of hard vs. soft X-ray
Histology	High quality images	Time consuming; resolves a lot of structure and may overestimate counts	Autotechnicon, Microtome, Microscope	See Natanson's studies; works well for blue sharks but not as reliable with mako vertebrae
Centrum-face shadow method	Easy	Requires some chemical treatment; may underestimate counts on large sharks as	Light, Microscope	See Semba's studies

alternating bands
are narrower

Table 2: Summary of the translation for band pairs and edges for each national delegation created during the 2017 ISC shark working group age and growth workshop. Wide denotes what type of edge in each analysis indicates having a partial year of growth, narrow denotes the edge type where a full year is counted. Band pair simply indicates the terminology used in each method to denote a band pair. Birth band treatment indicates weather the birth band is included in the count or is simply the point after which continuing begins.

Nation	Method	Wide	Narrow	Band Pair	Treatment of birth band
US	Hard x-ray	Translucent	Opaque	Wide/Narrow	Not counted
Japan	Shadowing	Concave	Convex	Concave/convex	Counted
Mexico	Transmitted light	Opaque	Translucent	Wide/Narrow	Not counted
Taiwan	Soft x-ray	Translucent	Opaque	Wide/Narrow	Not counted
New Zealand	Reflected light	Opaque	Translucent	Wide/Narrow	Not counted

Table 3: Parameter estimates for the Phase 4 model using data standardized using β_US and the two-band pair per year hypothesis.

Name	Sex	Data set	Mean	SD	2.5%	Median	97.5%
L_1			65.085	0.523	64.084	65.072	66.098
μ_{L_∞}	Male		226.192	15.616	197.608	225.728	260.805
μ_{L_∞}	Female		276.350	16.598	250.993	274.377	317.483
μ_{L_2}	Male		225.341	15.534	196.800	224.727	259.870
μ_{L_2}	Female		274.956	16.384	249.652	272.798	315.429
L_2	Male	JP	216.600	7.940	202.387	215.999	233.597
L_2	Male	MX	251.552	21.189	214.390	249.241	300.199
L_2	Male	TW	208.649	3.730	201.321	208.601	216.133
L_2	Male	US	219.985	20.272	184.916	217.974	266.367
L_2	Female	JP	265.328	10.540	246.205	264.546	287.687
L_2	Female	MX	282.430	19.473	251.896	280.032	326.524
L_2	Female	TW	254.416	4.527	245.753	254.367	263.224
L_2	Female	US	288.685	23.403	251.507	286.796	339.164
μ_k	Male		0.138	0.019	0.101	0.138	0.175
μ_k	Female		0.130	0.014	0.102	0.129	0.155
k	Male	JP	0.280	0.027	0.230	0.279	0.338
k	Male	MX	0.124	0.018	0.089	0.123	0.162
k	Male	TW	0.187	0.010	0.168	0.187	0.209
k	Male	US	0.205	0.040	0.135	0.201	0.295
k	Female	JP	0.176	0.016	0.147	0.175	0.208
k	Female	MX	0.098	0.011	0.075	0.099	0.119
k	Female	TW	0.133	0.006	0.121	0.133	0.146
k	Female	US	0.149	0.024	0.107	0.148	0.201
Mean CV_1	Male		0.130	0.008	0.115	0.129	0.146
Mean CV_1	Female		0.131	0.009	0.116	0.130	0.151
CV_1	Male	JP	0.186	0.012	0.163	0.185	0.211
CV_1	Male	MX	0.086	0.008	0.072	0.085	0.104
CV_1	Male	TW	0.122	0.011	0.103	0.121	0.145
CV_1	Male	US	0.125	0.027	0.077	0.123	0.184
CV_1	Female	JP	0.197	0.011	0.175	0.196	0.220
CV_1	Female	MX	0.082	0.007	0.068	0.082	0.097
CV_1	Female	TW	0.116	0.010	0.099	0.115	0.136
CV_1	Female	US	0.131	0.030	0.081	0.127	0.199
Mean CV_2	Male		0.065	0.008	0.052	0.064	0.082
Mean CV_2	Female		0.075	0.009	0.058	0.074	0.095
CV_2	Male	JP	0.052	0.008	0.036	0.051	0.070
CV_2	Male	MX	0.063	0.014	0.039	0.062	0.094
CV_2	Male	TW	0.045	0.006	0.034	0.044	0.058
CV_2	Male	US	0.101	0.028	0.058	0.099	0.164
CV_2	Female	JP	0.049	0.009	0.033	0.048	0.068
CV_2	Female	MX	0.070	0.016	0.044	0.069	0.108
CV_2	Female	TW	0.065	0.007	0.054	0.065	0.079
CV_2	Female	US	0.114	0.033	0.061	0.110	0.189

Table 4: Parameter estimates for the Phase 4 model using data standardized using β_{Japan} and the one-band pair per year hypothesis.

Name	Sex	Data set	Mean	SD	2.5%	Median	97.5%
L_1			65.123	0.520	64.088	65.109	66.105
μ_{L_∞}	Male		238.765	18.828	205.084	236.901	279.714
μ_{L_∞}	Female		307.778	21.548	270.644	305.524	356.002
μ_{L_2}	Male		236.912	18.430	203.837	235.197	276.632
μ_{L_2}	Female		302.440	20.617	266.535	299.979	348.914
L_2	Male	JP	216.832	8.933	200.634	216.386	236.319
L_2	Male	MX	279.267	28.959	230.389	276.237	340.978
L_2	Male	TW	225.747	5.673	214.699	225.865	236.750
L_2	Male	US	222.296	24.571	179.893	220.739	274.878
L_2	Female	JP	265.117	13.963	238.493	264.464	292.534
L_2	Female	MX	335.888	32.742	283.800	332.636	406.251
L_2	Female	TW	290.933	6.707	278.181	290.644	303.830
L_2	Female	US	312.947	32.365	259.501	307.857	385.491
μ_k	Male		0.119	0.017	0.087	0.119	0.153
μ_k	Female		0.102	0.018	0.069	0.101	0.137
k	Male	JP	0.199	0.022	0.159	0.198	0.243
k	Male	MX	0.057	0.013	0.033	0.056	0.084
k	Male	TW	0.096	0.007	0.083	0.095	0.110
k	Male	US	0.127	0.032	0.075	0.123	0.206
k	Female	JP	0.126	0.015	0.100	0.125	0.159
k	Female	MX	0.035	0.010	0.017	0.035	0.054
k	Female	TW	0.058	0.004	0.051	0.058	0.067
k	Female	US	0.084	0.020	0.048	0.083	0.125
Mean CV_1	Male		0.127	0.008	0.113	0.127	0.145
Mean CV_1	Female		0.131	0.009	0.116	0.130	0.151
CV_1	Male	JP	0.186	0.012	0.164	0.186	0.212
CV_1	Male	MX	0.083	0.007	0.069	0.082	0.099
CV_1	Male	TW	0.119	0.010	0.099	0.118	0.141
CV_1	Male	US	0.122	0.027	0.077	0.120	0.183
CV_1	Female	JP	0.197	0.011	0.177	0.197	0.221
CV_1	Female	MX	0.080	0.007	0.068	0.080	0.093
CV_1	Female	TW	0.116	0.008	0.102	0.116	0.132
CV_1	Female	US	0.131	0.030	0.078	0.129	0.200
Mean CV_2	Male		0.066	0.008	0.052	0.066	0.085
Mean CV_2	Female		0.073	0.010	0.055	0.071	0.094
CV_2	Male	JP	0.052	0.009	0.036	0.052	0.071
CV_2	Male	MX	0.068	0.015	0.042	0.067	0.100
CV_2	Male	TW	0.040	0.007	0.027	0.040	0.055
CV_2	Male	US	0.105	0.027	0.061	0.101	0.167
CV_2	Female	JP	0.051	0.010	0.034	0.050	0.071
CV_2	Female	MX	0.079	0.019	0.046	0.077	0.120
CV_2	Female	TW	0.045	0.007	0.033	0.045	0.060
CV_2	Female	US	0.116	0.034	0.063	0.111	0.190

Table 5: Log-likelihood by data component for each of the two Phase 4 models.

Data component	Scenario	Log-likelihood
Japan	β_{Japan} & 1 bp	611.0758
Japan	β_{US} & 2 bp	614.188
Mexico	β_{Japan} & 1 bp	505.5982
Mexico	β_{US} & 2 bp	506.8334
Taiwan	β_{Japan} & 1 bp	606.6466
Taiwan	β_{US} & 2 bp	594.4247
US	β_{Japan} & 1 bp	35.23311
US	β_{US} & 2 bp	35.68446

Table 6: Difference in expected log pointwise predictive density (ELPD) using leave-one-out cross-validation. The best model is shown in the top row.

Model	Difference (ELPD)	Standard error (ELPD)
Phase 4: β_{Japan} & 1 bp	0.00	0.00
Phase 4: β_{US} & 2 bp	-5.93	5.88
Phase 3: β_{Japan} & 1 bp	-5647.60	14.53
Phase 3: β_{US} & 2 bp	-5652.26	13.52

Table 7: Difference in expected log pointwise predictive density (ELPD) using leave-one-out cross-validation. The best model is shown in the top row.

Model	Difference (ELPD)	Standard error (ELPD)
Phase 4: β_{US} & 2 bp; no TW data	0.00	0.00
Phase 4: β_{Japan} & 1 bp; no TW data	-6.14	3.69

Table 8: Parameter estimates for the Phase 5 model using the Runcie et al. 2016 length frequency data and vertebral age data standardized using β_{US} and the two-band pair per year hypothesis.

Name	Sex	Data set	Mean	SD	2.5%	Median	97.5%
L_1			65.205	0.514	64.170	65.209	66.191
μ_{L_∞}	Male		226.140	14.137	200.788	224.864	256.596
μ_{L_∞}	Female		273.687	14.924	248.715	272.387	307.512
μ_{L_2}	Male		225.368	14.001	200.281	224.051	255.377
μ_{L_2}	Female		272.225	14.624	247.746	270.946	305.124
L_2	Male	JP	217.878	8.355	202.440	217.325	236.181
L_2	Male	MX	249.759	21.141	213.122	247.539	297.868
L_2	Male	TW	209.018	3.863	201.440	209.035	216.651
L_2	Male	US	221.057	19.446	187.369	218.988	266.026
L_2	Male	US – LF	227.789	25.250	182.613	225.353	286.649
L_2	Female	JP	266.289	10.854	246.494	266.022	289.184
L_2	Female	MX	280.433	18.619	250.016	278.478	322.078
L_2	Female	TW	254.426	4.832	245.318	254.434	264.221
L_2	Female	US	286.669	22.841	251.735	283.578	335.253
L_2	Female	US – LF	265.879	23.882	216.555	264.196	318.661
μ_k	Male		0.141	0.019	0.103	0.142	0.178
μ_k	Female		0.128	0.014	0.101	0.128	0.156
k	Male	JP	0.275	0.028	0.224	0.274	0.333
k	Male	MX	0.125	0.019	0.090	0.125	0.166
k	Male	TW	0.187	0.011	0.167	0.186	0.208
k	Male	US	0.202	0.037	0.137	0.200	0.283
k	Male	US – LF	0.145	0.032	0.089	0.143	0.217
k	Female	JP	0.174	0.016	0.145	0.174	0.207
k	Female	MX	0.099	0.011	0.077	0.100	0.121
k	Female	TW	0.133	0.007	0.120	0.133	0.146
k	Female	US	0.150	0.024	0.108	0.148	0.201
k	Female	US – LF	0.110	0.020	0.073	0.109	0.152
Mean CV_1	Male		0.129	0.009	0.113	0.128	0.148
Mean CV_1	Female		0.131	0.008	0.116	0.130	0.148
CV_1	Male	JP	0.185	0.012	0.163	0.185	0.211
CV_1	Male	MX	0.085	0.008	0.071	0.085	0.101
CV_1	Male	TW	0.121	0.011	0.101	0.121	0.146
CV_1	Male	US	0.123	0.028	0.076	0.120	0.190
CV_1	Male	US – LF	0.101	0.030	0.056	0.097	0.169
CV_1	Female	JP	0.195	0.011	0.174	0.195	0.218
CV_1	Female	MX	0.082	0.007	0.069	0.082	0.097
CV_1	Female	TW	0.115	0.009	0.098	0.115	0.133
CV_1	Female	US	0.130	0.028	0.082	0.128	0.193
CV_1	Female	US – LF	0.097	0.028	0.052	0.094	0.163
Mean CV_2	Male		0.066	0.009	0.051	0.065	0.086
Mean CV_2	Female		0.075	0.010	0.059	0.074	0.096
CV_2	Male	JP	0.051	0.010	0.034	0.051	0.071
CV_2	Male	MX	0.063	0.014	0.039	0.062	0.094
CV_2	Male	TW	0.045	0.006	0.033	0.045	0.058
CV_2	Male	US	0.103	0.030	0.057	0.099	0.175
CV_2	Male	US – LF	0.110	0.035	0.057	0.104	0.191
CV_2	Female	JP	0.048	0.009	0.033	0.048	0.067
CV_2	Female	MX	0.071	0.016	0.043	0.069	0.106
CV_2	Female	TW	0.065	0.007	0.053	0.065	0.079
CV_2	Female	US	0.115	0.033	0.062	0.111	0.194
CV_2	Female	US – LF	0.108	0.031	0.059	0.104	0.181

Table 9: Parameter estimates for the mixture modeling component of the Phase 5 model using the Runcie et al. 2016 length frequency data and vertebral age data standardized using β_{US} and the two-band pair per year hypothesis.

Name	Sex	Mode	Mean	SD	2.5%	Median	97.5%
θ	Male	1	0.070	0.020	0.019	0.073	0.105
θ	Male	2	0.509	0.043	0.449	0.503	0.629
θ	Male	3	0.349	0.030	0.285	0.349	0.405
θ	Male	4	0.072	0.017	0.035	0.073	0.102
θ	Female	1	0.109	0.019	0.073	0.109	0.146
θ	Female	2	0.573	0.028	0.519	0.571	0.629
θ	Female	3	0.261	0.025	0.214	0.261	0.310
θ	Female	4	0.057	0.012	0.036	0.056	0.083
CV_{mix}	Male		0.105	0.015	0.088	0.102	0.148
CV_{mix}	Female		0.092	0.007	0.081	0.092	0.107
a_0	Male		0.294	0.150	0.031	0.286	0.613
a_0	Female		0.345	0.160	0.065	0.344	0.660
L_{obs}	Male	1	67.560	1.928	63.622	67.672	70.701
L_{obs}	Male	2	90.889	1.062	88.957	90.828	93.141
L_{obs}	Male	3	117.819	1.938	114.810	117.617	122.118
L_{obs}	Male	4	157.264	4.443	150.791	156.927	166.269
L_{obs}	Female	1	69.243	1.238	66.933	69.227	71.616
L_{obs}	Female	2	91.439	0.776	89.862	91.454	92.871
L_{obs}	Female	3	118.882	1.569	115.770	118.925	121.923
L_{obs}	Female	4	157.832	3.425	150.863	157.829	164.447
L_{true}	Male	1	68.350	3.249	62.177	68.272	74.911
L_{true}	Male	2	90.008	4.337	81.883	89.770	98.969
L_{true}	Male	3	113.565	5.624	102.616	113.392	124.491
L_{true}	Male	4	145.930	8.075	129.927	145.944	161.958
L_{true}	Female	1	70.276	3.178	64.445	70.134	76.739
L_{true}	Female	2	91.450	4.067	83.793	91.288	99.793
L_{true}	Female	3	115.166	5.227	105.120	115.182	125.985
L_{true}	Female	4	144.448	7.466	129.741	144.360	159.536

Figures

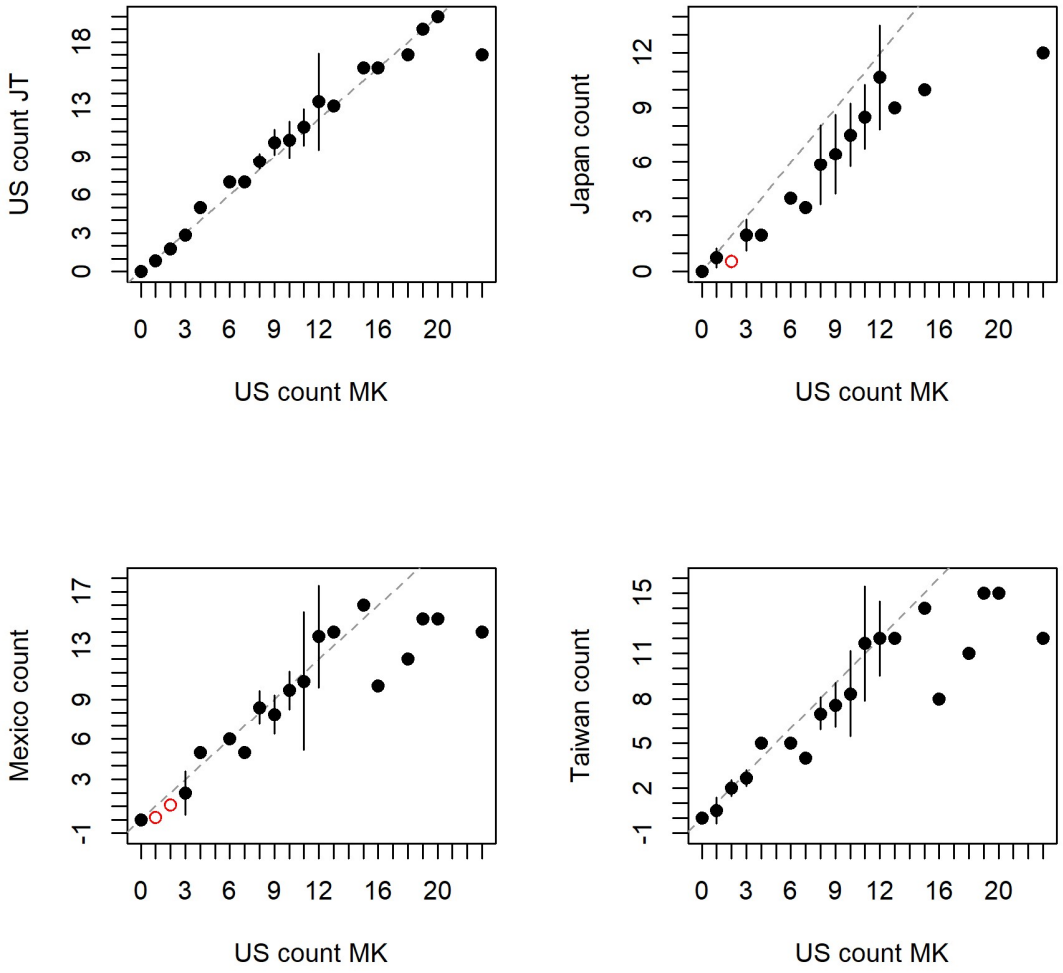


Figure 1: Comparison plots of country specific band pair counts of ISC mako vertebra reference collection.

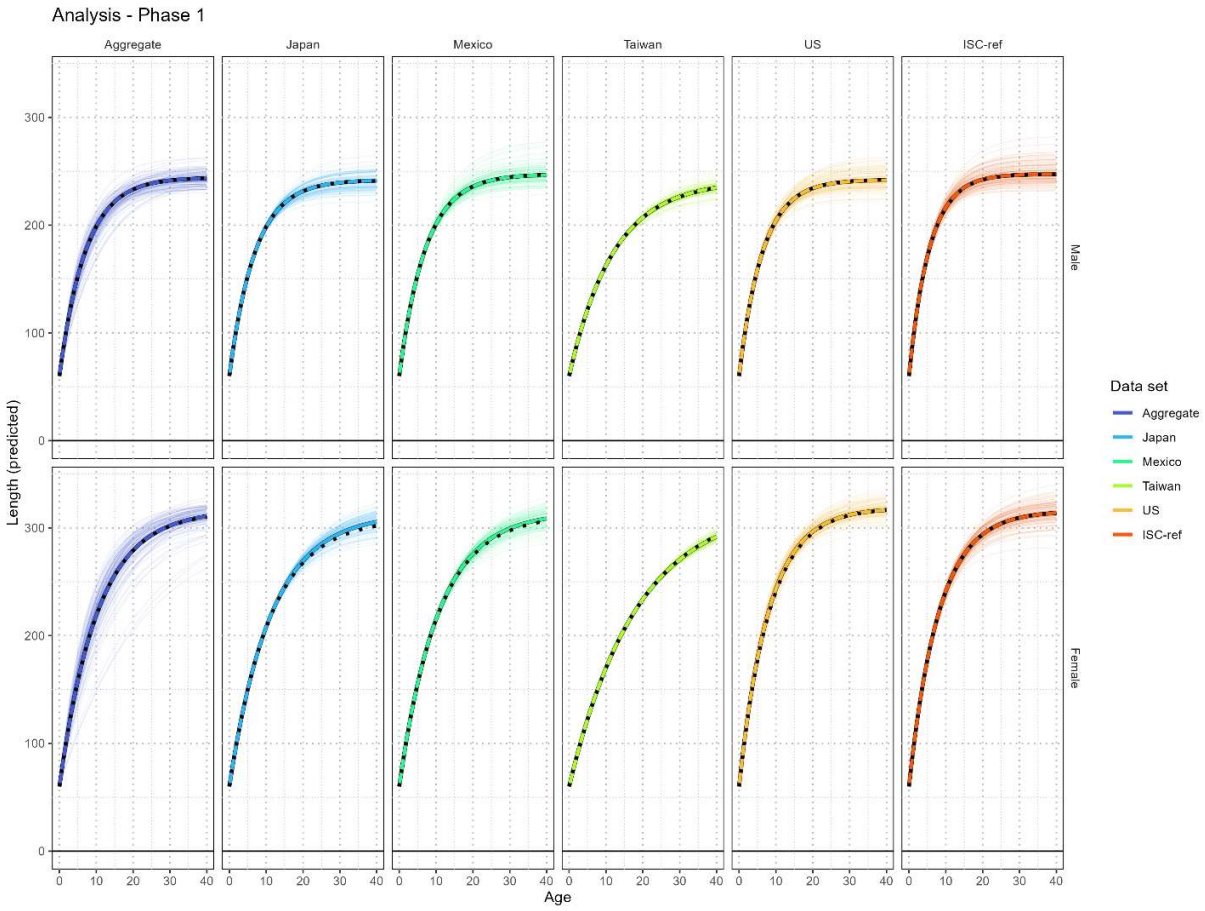


Figure 2: The growth curves given by the hyperdistribution (Aggregate) and the 5 data-set specific random effect growth curves for the Analysis – Phase 1 model. The narrow transparent lines are individual draws from the posterior distribution, the thicker line with the dark outline is the median curve in each panel. The black dotted line is the equivalent median curve from Takahashi et al. 2017 Table A3.

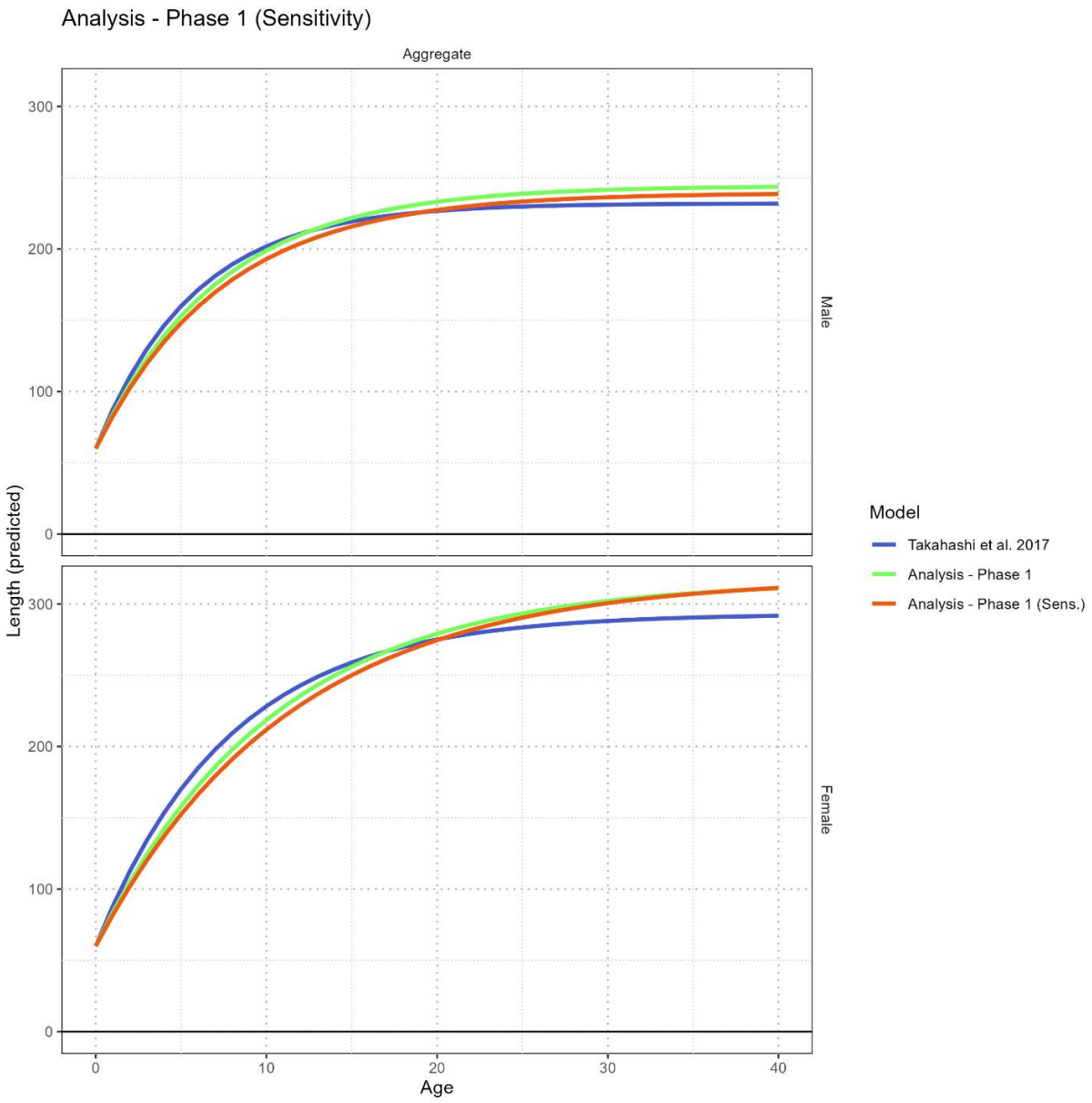


Figure 3: A comparison of the Takahashi et al. 2017 model showing the sensitivity to the included data sets. The blue line is the Takahashi et al. 2017 model fit to all 7 data-sets (Takahashi et al. 2017; Table A1). The green line is the Takahashi et al. 2017 model fit to the 4 age-length data sets and the ISC reference collection (Takahashi et al. 2017; Table A3). The orange line (this analysis) fits the same model only to the 4 age-length data sets.

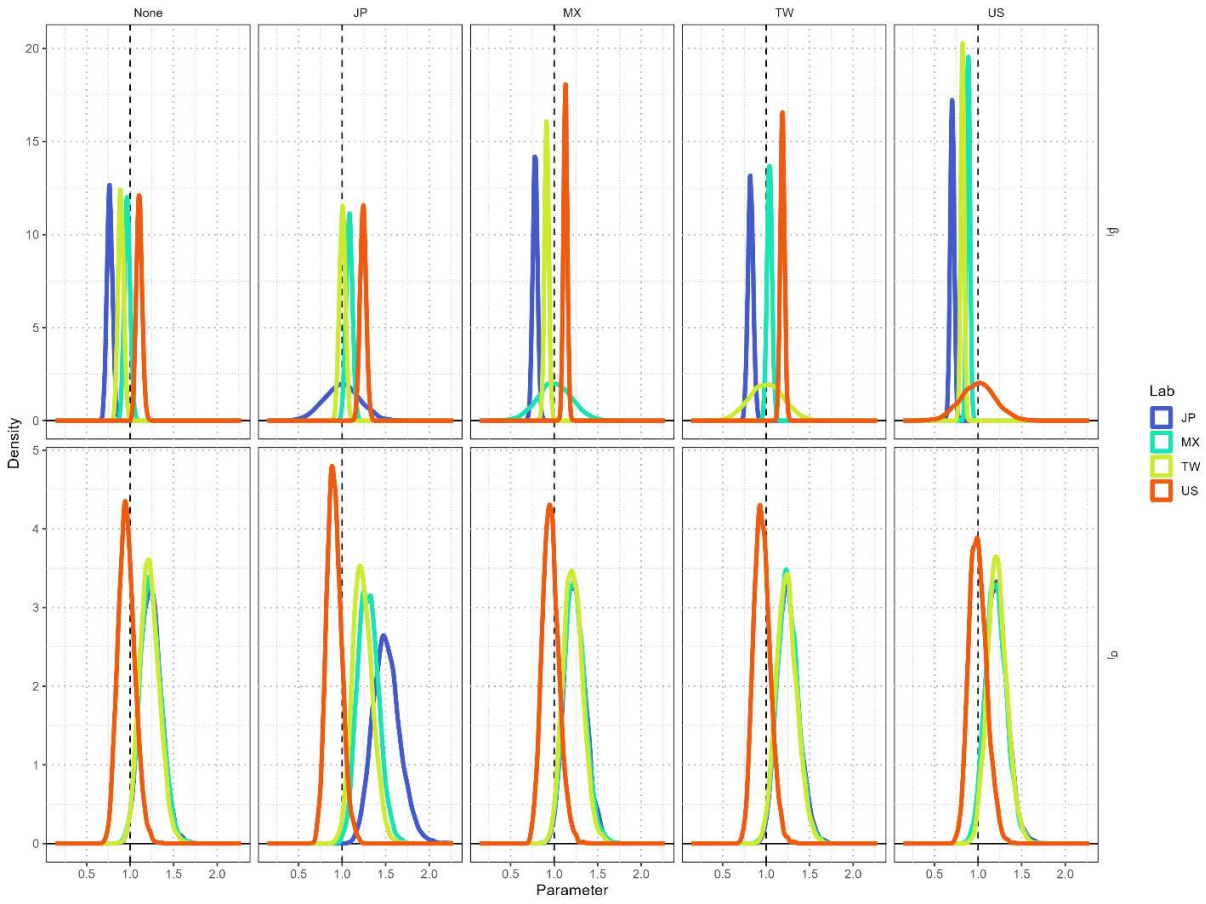


Figure 4: Estimates of lab calibration factor β_j and observation error σ_j by lab (colors) and across scenarios where either no lab is assumed to have the true methodology (None) or each lab in turn is assumed to have the true methodology.

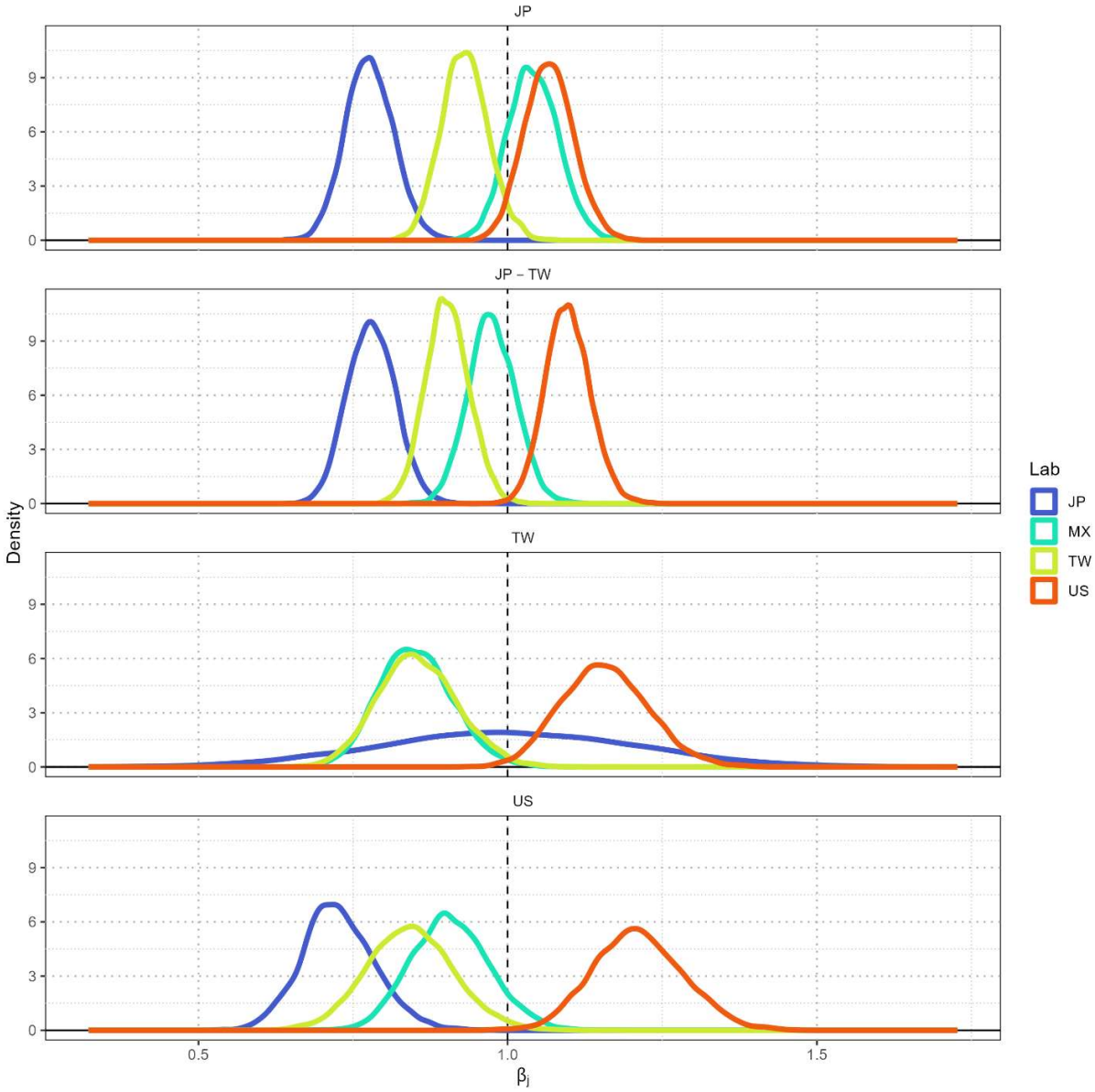


Figure 5: Estimates of lab calibration factor β_j by data partition (e.g., Japan samples only, Taiwan samples only, Japan & Taiwan samples, and US samples only) where no lab is assumed to have the true methodology.

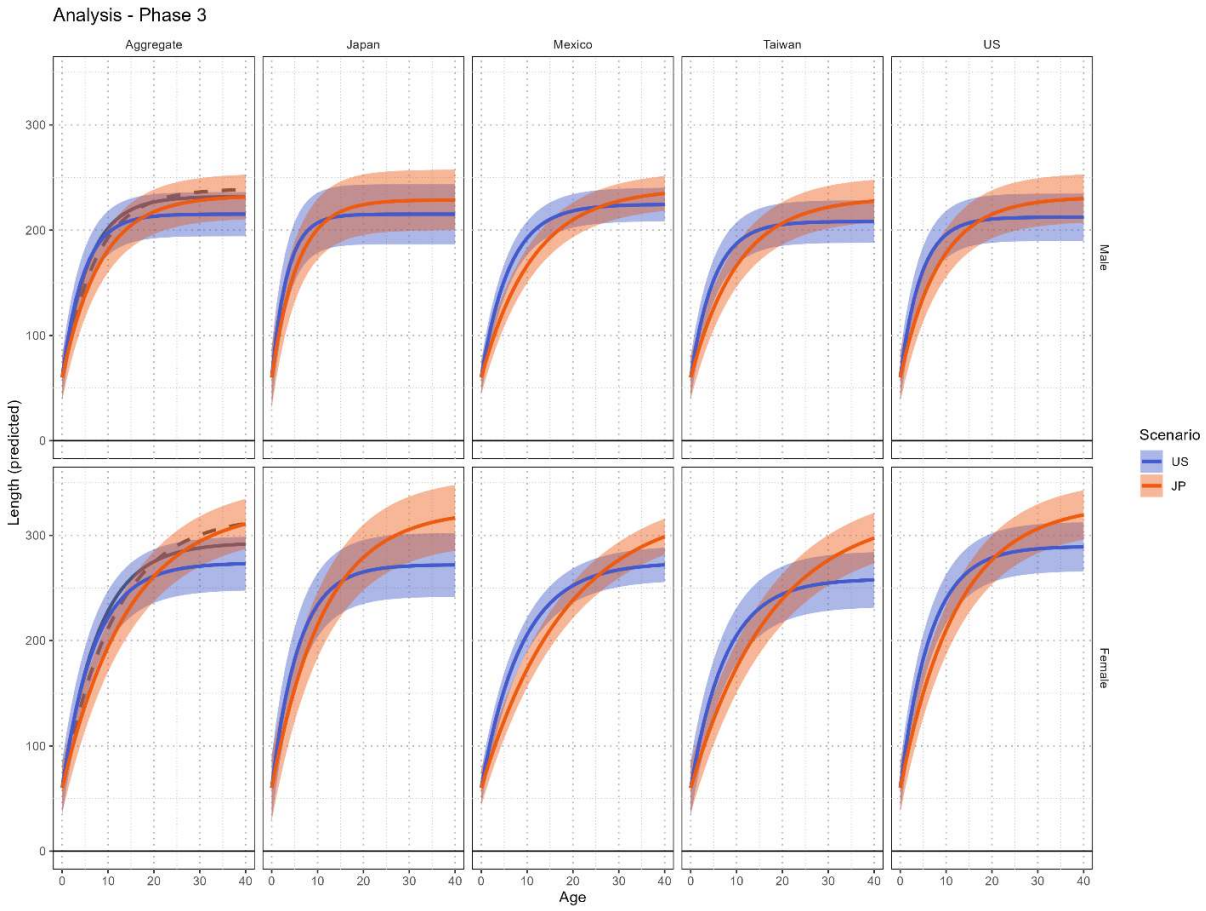


Figure 6: The growth curves given by the hyperdistribution (Aggregate) and the 4 data-set specific random effect growth curves for the Analysis – Phase 3 model. The two alternative hypotheses are shown by the colors: β_{US} and the two-band pair per year hypothesis (blue), and β_{Japan} and the one-band pair per year hypothesis (orange). The Takahashi et al. 2017 growth curve used in the previous assessment is shown by the solid gray line, and the gray dotted line gives the equivalent curve fit to the 4 unstandardized data-sets.

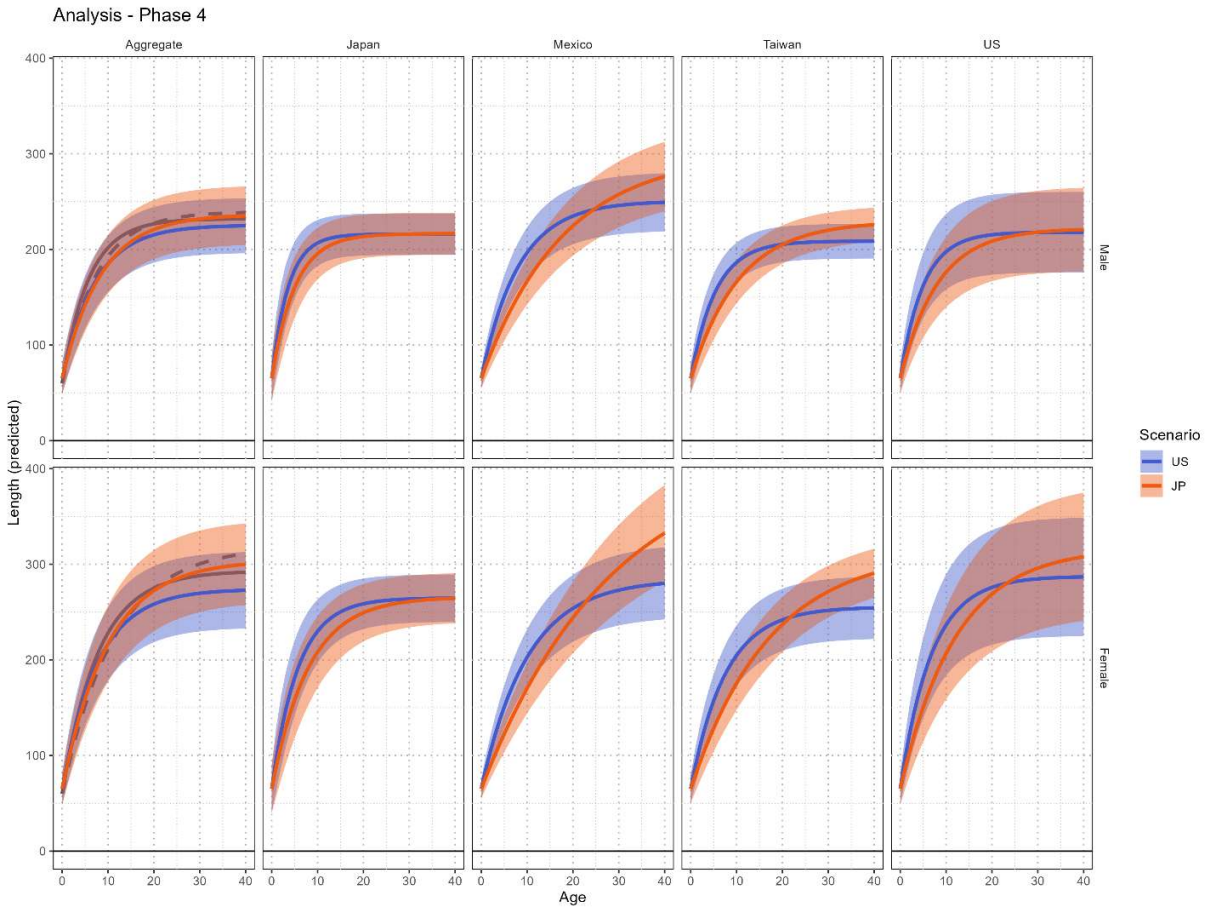


Figure 7: The growth curves given by the hyperdistribution (Aggregate) and the 4 data-set specific random effect growth curves for the Analysis – Phase 4 model. The two alternative hypotheses are shown by the colors: β_{US} and the two-band pair per year hypothesis (blue), and β_{Japan} and the one-band pair per year hypothesis (orange). The Takahashi et al. 2017 growth curve used in the previous assessment is shown by the solid gray line, and the gray dotted line gives the equivalent curve fit to the 4 unstandardized data-sets.

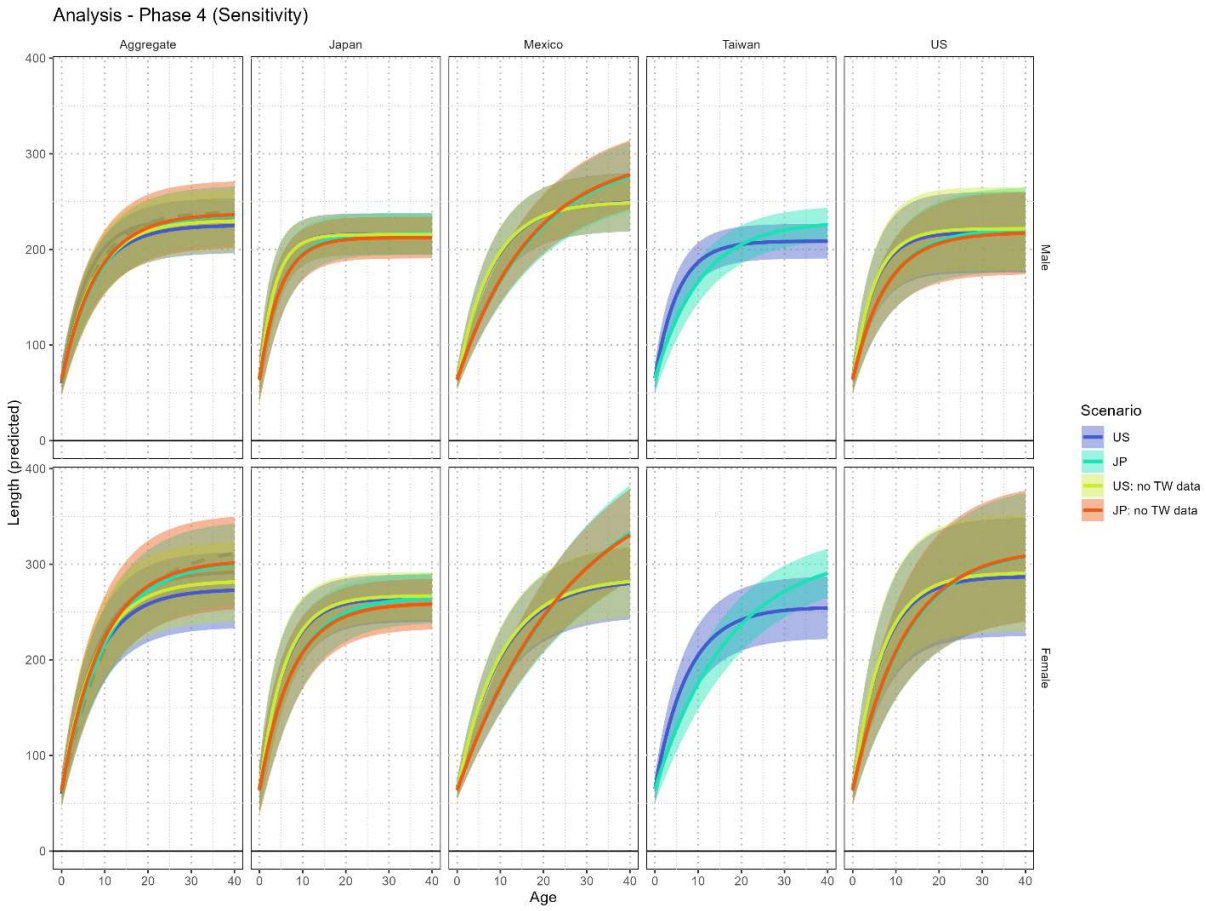


Figure 8: The growth curves given by the hyperdistribution (Aggregate) and the 4 data-set specific random effect growth curves for the Analysis – Phase 4 model. The two 4 scenarios shown include the following alternative hypotheses with and without the inclusion of the Taiwanese data: β_{US} and the two-band pair per year hypothesis (blue), and β_{Japan} and the one-band pair per year hypothesis (orange). The Takahashi et al. 2017 growth curve used in the previous assessment is shown by the solid gray line, and the gray dotted line gives the equivalent curve fit to the 4 unstandardized data-sets.

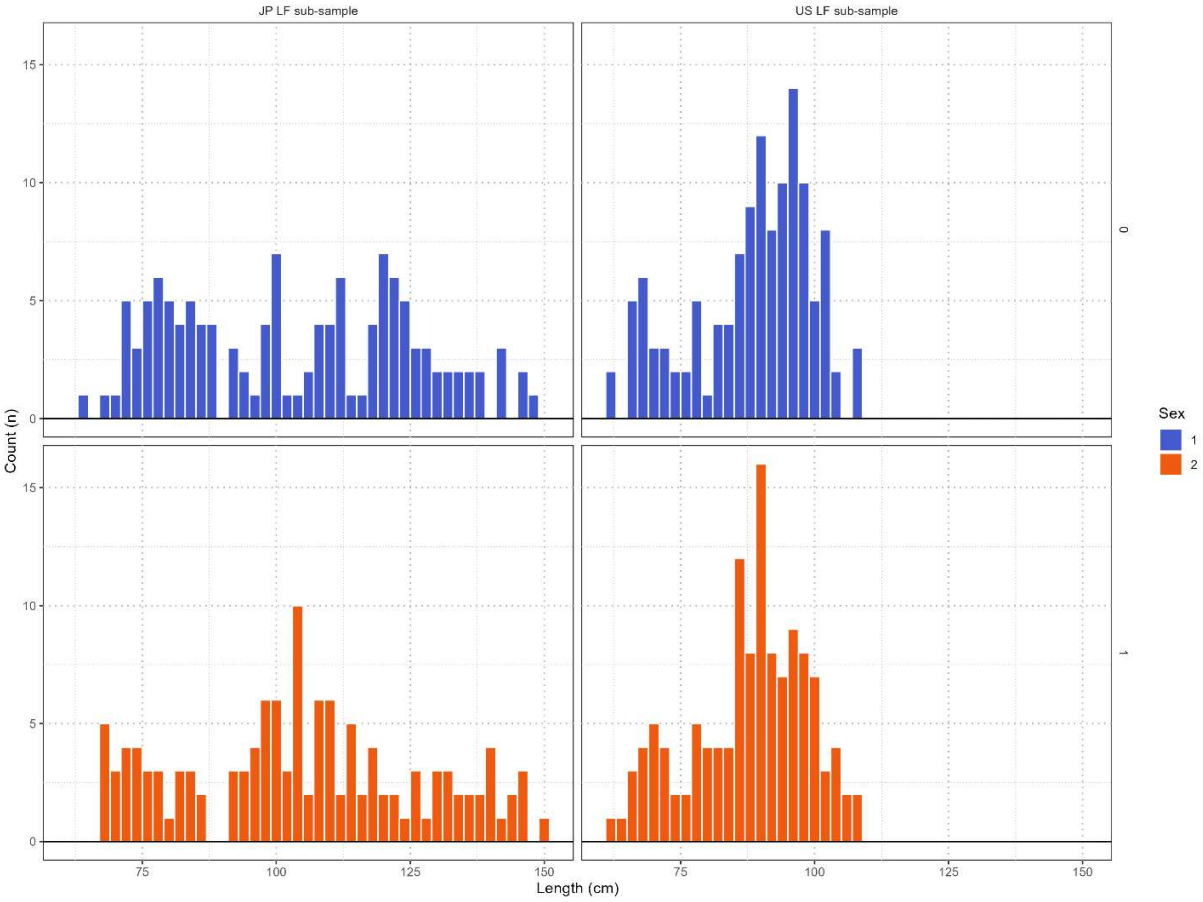


Figure 9: Subsets of length frequency data from Japan and the US from Takahashi et al. (2017). Males are shown in blue and females are shown in orange.

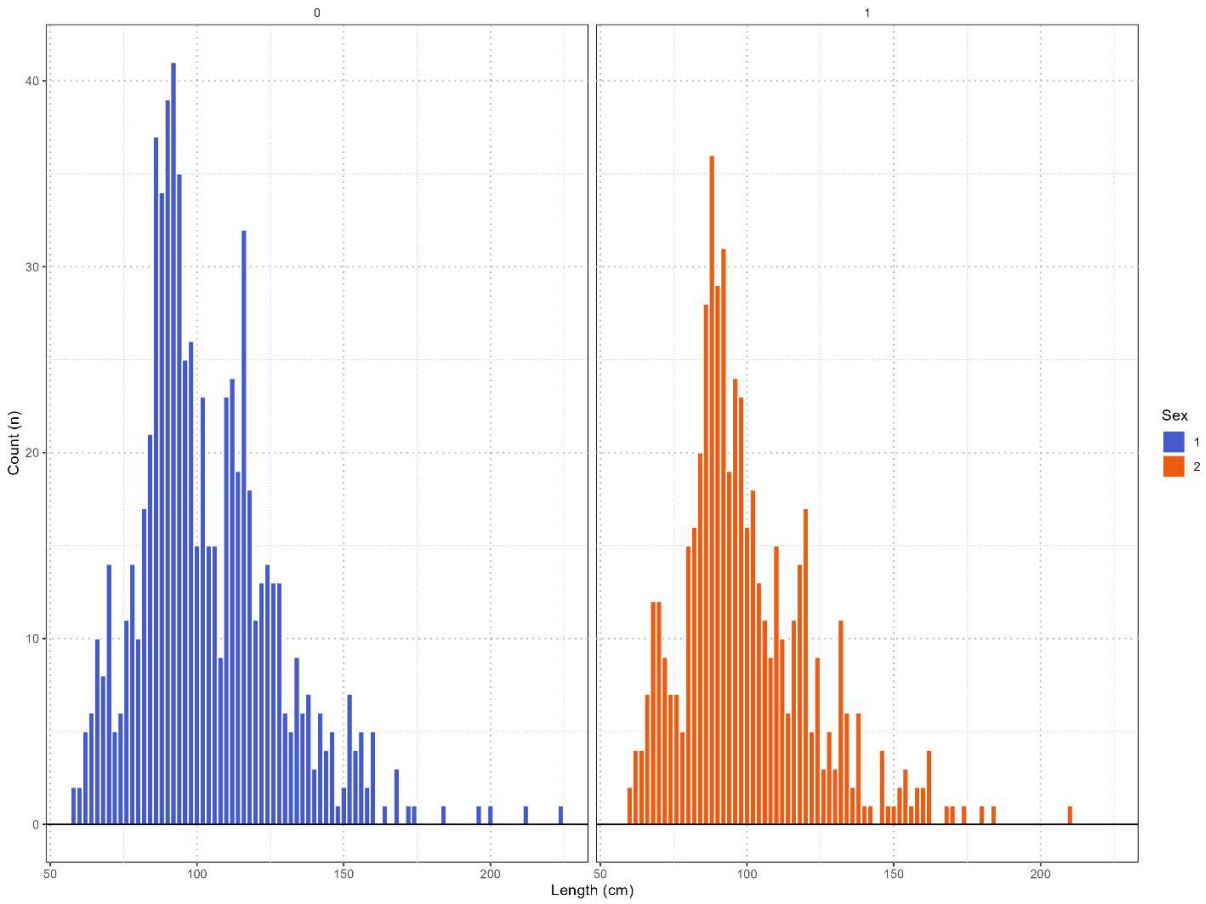


Figure 10: The full length composition data from all individuals with identified sex from Runcie et al. 2016. Males are shown in blue and females are shown in orange.

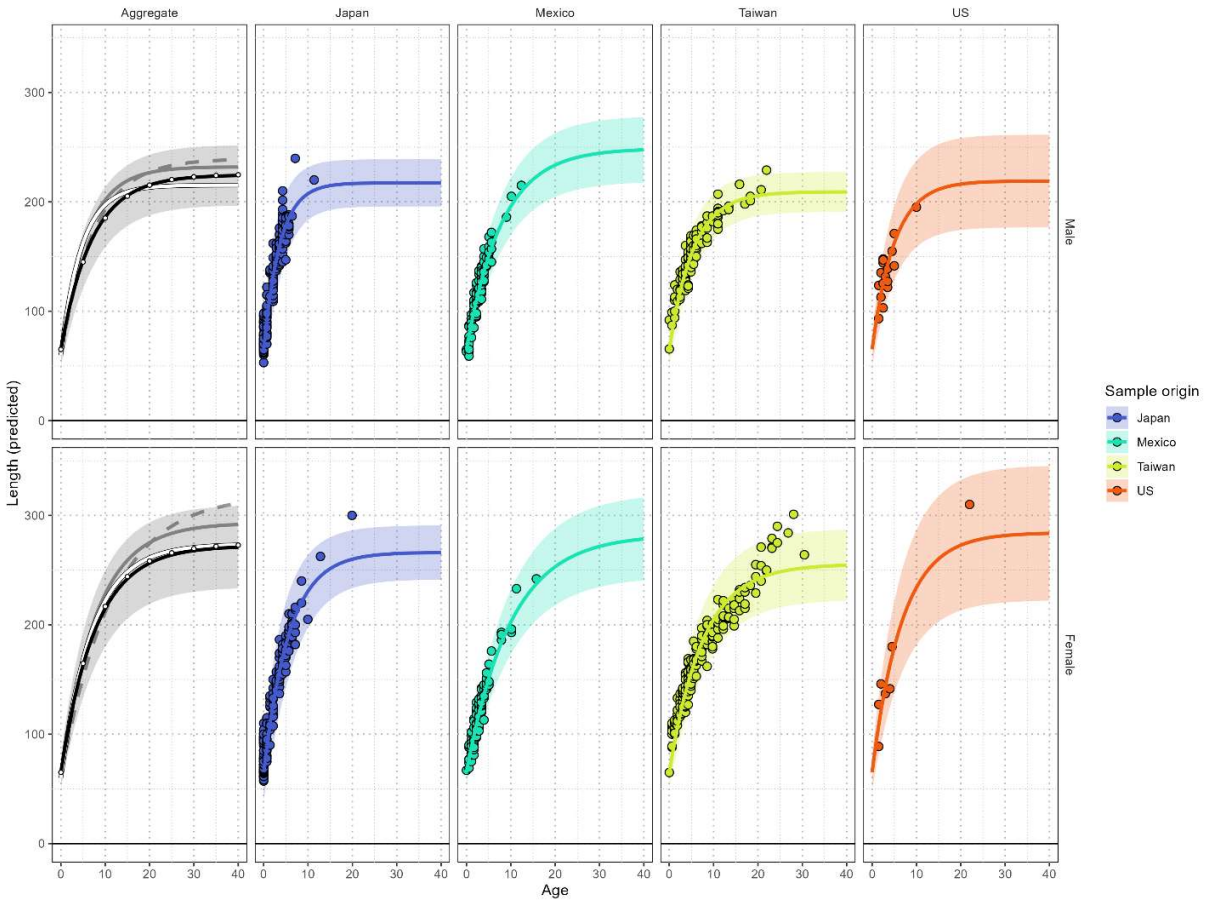


Figure 11: The growth curves given by the hyperdistribution (Aggregate) and the 4 data-set specific random effect growth curves for the Analysis – Phase 5 model. The corresponding Phase 4 model is shown in the Aggregate panel using white dots. The corresponding Phase 3 model is shown in the Aggregate panel using a solid white line. The Takahashi et al. 2017 growth curve used in the previous assessment is shown by the solid gray line, and the gray dotted line gives the equivalent curve fit to the 4 unstandardized data-sets.

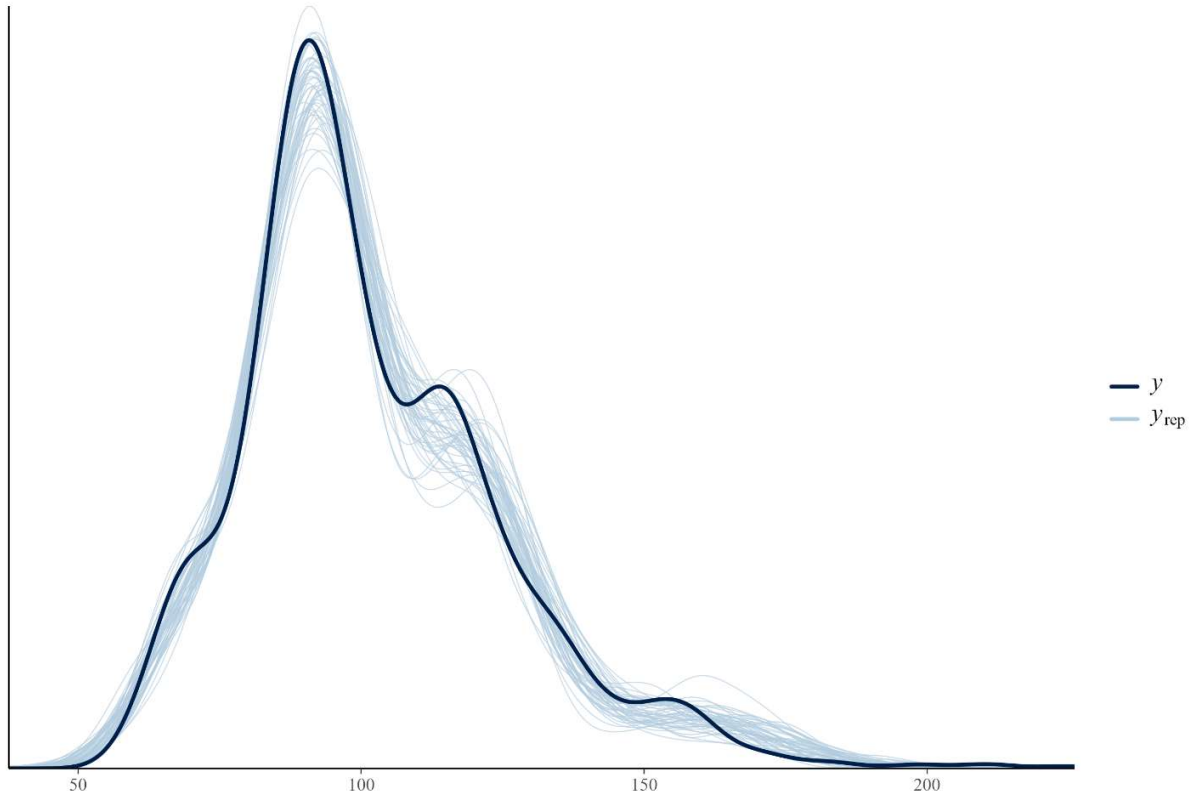


Figure 12: Distribution of Runcie et al. 2016 length frequency observations fit within Analysis – Phase 5 (y) relative to simulated posterior predicted distributions of the lengths (yrep).

Cailliet, Gregor M, and DW Bedford. 1983. 'The biology of three pelagic sharks from California waters, and their emerging fisheries: a review', *CalCOFI Rep*, 24: 57-69.
Hsu, Hua-Hsun. 2003. 'Age, growth, and reproduction of shortfin mako, Isurus oxyrinchus in the northwestern Pacific', MS thesis, National Taiwan Ocean Univ., Keelung, Taiwan.
ISC. 2011. "Report of the first shark age and growth workshop." In *Report of the Meeting of the International Scientific Committee for Tuna and Tuna-like Species in the North Pacific Ocean.* La Jolla, California.
———. 2014. "Report of the second shark age and growth workshop." In *Report of the Meeting of the International Scientific Committee for Tuna and Tuna-like Species in the North Pacific Ocean.* La Jolla, California.
———. 2015. "Indicator-based analysis of the status of shortfin mako shark in the North Pacific Ocean." In *Report of the Fifteenth Meeting of the International Scientific Committee for Tuna and Tuna-like Species in the North Pacific Ocean.*, 76. Kona, Hawaii.
———. 2017a. "Report of the ISC Shark WG data prep meeting on mako sharks " In *Report of the 18th Meeting of the International Scientific Committee for Tuna and Tuna-like Species in the North Pacific Ocean.*, 76. Shimizu, Shizoka, Japan.
———. 2017b. "Report of the third shark age and growth workshop." In *Report of the Meeting of the International Scientific Committee for Tuna and Tuna-like Species in the North Pacific Ocean.* Shimizu, Shizoka, Japan.
———. 2018. "Stock Assessment of Shortfin Mako Shark in the North Pacific Ocean Through 2016." In.
Joung, Shoou-Jeng, and Hua-Hsun Hsu. 2005. 'Reproduction and embryonic development of the shortfin mako, *Isurus oxyrinchus* Rafinesque, 1810, in the northwestern Pacific', *Zoological Studies-Taipei.*, 44: 487-96.
Kai, M, K Shiozaki, S Ohshimo, and K Yokawa. 2015. 'Growth and spatiotemporal distribution of juvenile shortfin mako (*Isurus oxyrinchus*) in the western and central North Pacific', *Marine and Freshwater Research*.
Kinney, MJ, RJD Wells, and S Kohin. 2016. 'Oxytetracycline age validation of an adult shortfin mako shark Isurus oxyrinchus after 6 years at liberty', *Journal of Fish Biology*.
Okamura, Hiroshi, and Yasuko Semba. 2009. 'A novel statistical method for validating the periodicity of vertebral growth band formation in elasmobranch fishes', *Canadian Journal of Fisheries and Aquatic Sciences*, 66: 771-80.
R Core Team. 2024. "R: A language and environment for statistical computing. R Foundation for Statistical Computing." In. Vienna, Austria.
Ribot-Carballal, MC, Felipe Galván-Magaña, and Casimiro Quiñónez-Velázquez. 2005. 'Age and growth of the shortfin mako shark, Isurus oxyrinchus, from the western coast of Baja California Sur, Mexico', *Fisheries Research*, 76: 14-21.
Runcie, Rosa, David Holts, James Wraith, Yi Xu, Darlene Ramon, Rand Rasmussen, and Suzanne Kohin. 2016. 'A fishery-independent survey of juvenile shortfin mako (*Isurus oxyrinchus*) and blue (*Prionace glauca*) sharks in the Southern California Bight, 1994–2013', *Fisheries Research*, 183: 233-43.
Semba, Yasuko, Ichiro Aoki, and Kotaro Yokawa. 2011. 'Size at maturity and reproductive traits of shortfin mako, *Isurus oxyrinchus*, in the western and central North Pacific', *Marine and Freshwater Research*, 62: 20-29.
Semba, Yasuko, Hideki Nakano, and Ichiro Aoki. 2009. 'Age and growth analysis of the shortfin mako, *Isurus oxyrinchus*, in the western and central North Pacific Ocean', *Environmental Biology of Fishes*, 84: 377-91.

Stan Development Team. 2024a. "RStan: the R interface to Stan." In. https://mc-stan.org/.
———. 2024b. "Stan Modeling Language Users Guide and Reference Manual." In. https://mc-stan.org.
Takahashi, M., M. Kai, Y. Semba, M. Kanaiwa, K.M Liu, J.A. Rodríguez-Madrigal, J.T. Ávila, M.J. Kinney, and J.N. Taylor. 2017. "Meta-analysis of growth curve for shortfin mako shark in the North Pacific." In *ISC/17/SHARKWG-1/05*.
Wells, David RJ, Susan E Smith, Suzanne Kohin, Ellen Freund, Natalie Spear, and Darlene A Ramon. 2013. 'Age validation of juvenile Shortfin Mako (*Isurus oxyrinchus*) tagged and marked with oxytetracycline off southern California', *Fishery Bulletin*, 111: 147-60.
Zhou, Shijie, Sarah Martin, Dan Fu, and Rishi Sharma. 2019. 'A Bayesian hierarchical approach to estimate growth parameters from length data of narrow spread', *ICES Journal of Marine Science*, 77: 613-23.