**Supplementary material:** Using age compositions derived from spatio-temporal models and acoustic data collected by uncrewed surface vehicles to estimate Pacific Hake (*Merluccius productus*) biomass-at-age

Derek G. Bolser\*1,2, Aaron M. Berger3, Dezhang Chu4, Steve de Blois4, John Pohl4, Rebecca E. Thomas4, John Wallace4, Jim Hastie4, Julia Clemons3, Lorenzo Ciannelli1,5

1Cooperative Institute for Marine Ecosystem and Resources Studies, Oregon State University, Newport, OR, USA

2Present address: Assessment and Monitoring Division, Office of Science and Technology, National Marine Fisheries Service, National Oceanographic and Atmospheric Administration, Silver Spring, MD, USA

3Fishery Resource Analysis and Monitoring Division, Northwest Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, Newport, OR, USA

4Fishery Resource Analysis and Monitoring Division, Northwest Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, Seattle, WA, USA

5 College of Earth, Ocean, and Atmospheric Studies, Oregon State University, Corvallis, OR, USA

\*Corresponding author: Derek Bolser, [derek.bolser@noaa.gov](mailto:derek.bolser@noaa.gov)

*Supplementary material S1: Additional information on acoustic sampling and data processing protocols*

*Echogram scrutiny*

Species classification in the hake survey is based on the expert judgment of analysts and subsequently verified or refuted with midwater trawling. In Saildrone data, species classification was based on the judgment of analysts alone. Each transect is judged by at least two reviewers.  In general, analysts make judgements based on five types of information:

1. Acoustic properties attributes related to frequency response, e.g. the relative intensity of the signal at various acoustic frequencies. For swimbladder-bearing fish, such as hake, the frequency response is, in theory, the highest at the lowest survey frequency (I.e., 18 kHz for acoustic survey and 38 kHz for Saildrone survey application.
2. Morphometric attributes: mark intensity and variability, mark shape and size (length, thickness, wiggliness), aggregation patchiness. This is where expert judgment is most highly applied, as morhphology of hake schools can exhibit considerable variation.
3. Positional attributes: Depth of the mark, distance off bottom, distance from shelf and relations to other bathymetric features. Hake are typically located between 75 and 500 m depth in the water column, and are often associated with the shelf break area, although are also found both on the shelf and offshore.
4. Ancillary attributes: distance from other marks and aggregations, time of day, other ancillary information. For example, hake form midwater shoals daytime in summer months, and in the other seasons may remain in a diffuse shoal at night instead of dispersing.
5. Historical, regional, and local contextual attributes.  For example, hake in the northern portion of their range can be easily mistaken for Walleye pollock and vice versa, leading to an additional need to trawl on echosign that looks like Walleye pollock.

Since analysts had less acoustic frequencies to work with in Saildrone data, classifications were naturally more uncertain than in the acoustic trawl survey data. Further, without verification trawls to validate species classification, analysts had to designate aggregations based on acoustic data alone. Due to many inherent differences between the regular ship-based acoustics and the Saildrones, the acoustic data collected in Saildrone surveys may have significant differences despite being judged in as similar a manner possible to the acoustic trawl surveys.

*Mixed species data processing*

If ATS verification trawls indicated that an acoustic region was mixed-species in nature, additional steps are necessary to estimate the NASC of hake in the aggregation. In Saildrone sampling, aggregations were designated as either 100% or 0% hake based on the judgement of analysts as trawling for verification was not possible. The steps to estimate hake NASC from a mixed species aggregation the following:

1. Generate a length-weight relationship for non-hake species in the trawl
2. Estimate backscattering cross section of hake using the Traynor (1996) target strength-length model
3. Estimate the backscattering cross section of non-hake species using published target strength-length models
4. Calculating total backscattering contribution of each species group in the haul by multiplying numbers of individuals by backscattering cross section
5. Calculate a hake:non-hake NASC ratio, then scale from haul area to aggregation area to calculate total NASC for each species.

*Supplementary material S2: Total biomass estimate sensitivity analysis for the impact of stratum 0 predictions*

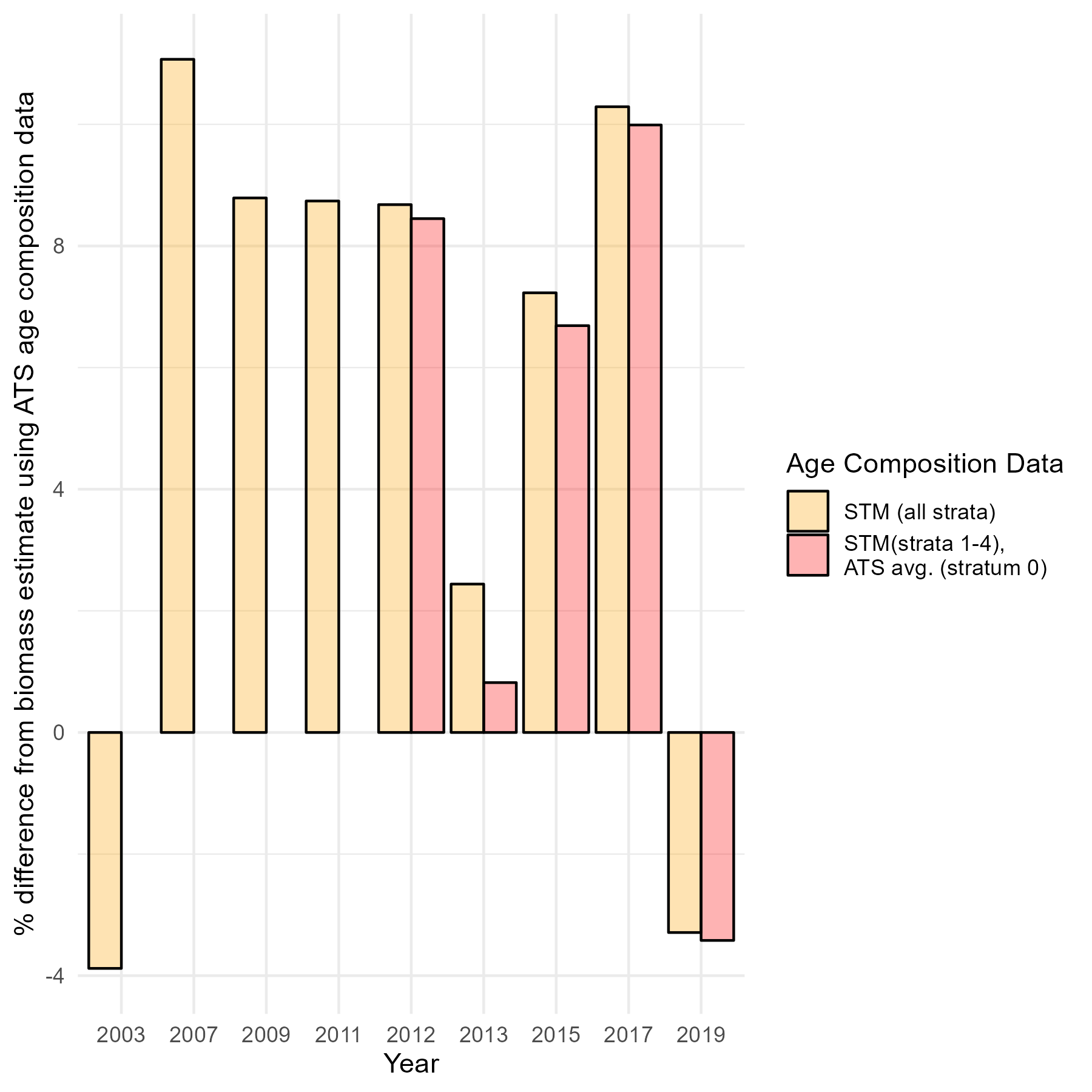


Fig. S2.1. Percent difference in age 2+ hake biomass estimates derived from different sources of age composition data using ATS acoustics. Age compositions include comparing observations from the ATS with STM (all years; yellow), and the ATS with STM in strata 1-4 combined with the ATS average age compositions for stratum 0 due to a lack of data in that location (2012-2019; red). Although biological sampling was conducted in stratum 0 in 2009 and 2011, hake older than age two were not detected in acoustic data, so it was not possible to include stratum 0 data directly in biomass estimates. Accordingly, the age class composition data source shown in red was only used in biomass estimates from 2012-2019.

With the exception of 2019, estimating biomass using STM age composition estimates, with the addition of average composition from the ATS for stratum 0 (exclusive of the year that biomass was being estimated in), produced biomass estimates that were marginally more similar to those derived from ATS age composition data in that year than if using STM estimates alone for all strata.

Table S2.1 Biomass estimates (kilotonnes; kt) derived from different sources of data for 2019.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | ATS age composition data | STM (strata 1-4) & ATS average (stratum 0) age composition data | STM (all strata) age composition data | Difference between biomass estimates (age composition data difference: ATS & STM [strata 1-4] & ATS average [stratum 0]) | Difference between biomass estimates (age composition data difference: ATS & STM) |
| ATS - acoustic data | 1523.59 kt | 1471.45 kt | 1473.91 kt | -3.4% | -3.2% |
| USV - acoustic data | 1114.88 kt | 1118.51 tmtkt | 1143.80 kt | +0.3% | +2.6% |
| Difference between biomass estimates (acoustic data difference) | -26.8% | -24.0% | -22.4% |  |  |

*Supplementary material S3: Biomass-at-age estimates derived from different combinations of data*

C:\Users\Derek.Bolser\Documents\Saildrone postdoc\Manuscripts\Frontiers submission 2\Supp. Fig. biomass at age_revised_2.tiff

Fig. S3. Biomass at age estimates (kilotonnes; kt) derived from different combinations of data, including biomass at age estimates from the sensitivity analysis for the impact of STM predictions in INPFC stratum 0. Estimates from 2003-2017 employed ATS acoustic data and the legend reflects the source of age composition data (ATS, STM, or STM (INPFC strata 1-4) & ATS average (INPFC stratum 0)). The legend for 2019 indicated the source of age composition data (first; ATS, STM, STM (INPFC strata 1-4) & ATS average (INPFC stratum 0)) and the source of acoustic data (second; ATS or USV). Age 20 encompasses all hake over 20 years old.

*Supplementary material S4: Statistical evaluation of STM biomass-at-age class estimates for 2019.*

*Model evaluation methods*

In simulation testing, new fixed and random effects were simulated from the joint precision matrix (i.e., matrix of implied model parameter distributions) of the operating model, and these effects were used to simulate new data. This was accomplished using the *simulate()* function in *TMB* (ver. 1.7.20; Kristensen *et al*., 2016) and specifying a ‘type 3’ simulator. This process was replicated ten times and the model was fit to each of the new datasets. A regression was fit between the ‘true’ and ‘estimated’ biomass values for each age class at each time period, with the slope of the regression line (i.e., ‘hyperstability’ or ‘sensitivity’) indicating how well the model is able to make predictions at different levels of biomass. In R, the regression was specified as follows:

lm( log(Est) ~ 0 + factor(Replicate)+ log(True), data=x )

In a perfect scenario, a model would have a hyperstability value of one. The relative error of the estimation models was also calculated, which indicated the magnitude of error relative to the magnitude of the estimate. In a perfect scenario, the relative error of the model would be zero.

In k-fold cross validation, data were split into ten partitions and the model was fit to each partition. The likelihood of the model was recorded and the average likelihood across error-free fits was multiplied by ten (as models did not converge when fit to some partitions of data) and compared with the likelihood of the model fit to the full dataset to calculate a likelihood ratio. In a perfect scenario, the likelihood ratio of the model would be one.

*Model evaluation results*

The single-factor model configuration had a mean relative error of 1.55 and a median relative error of 1.37 in simulation testing. These numbers indicate that the magnitude of error between ‘true’ and ‘estimated’ biomass was over 100% of biomass. It was not possible to calculate hyperstability for this model due to convergence issues. The multi-factor model configuration had a mean relative error of 0.90 and a median relative error of 0.66, reflecting a magnitude of error that was less than 100% of biomass but still substantial. The hyperstability of the multi-factor model was 0.30, indicating that model predictions exhibited a low degree of sensitivity to variation in ‘true’ biomass. The mean relative error for the IID configured model was 1.10 and the median relative error was 0.52, indicating high variability in the magnitude of error across model runs, and a high degree of error overall. The hyperstability of the IID model was 0.13, which showed insensitivity of model predictions to variation in ‘true’ biomass. The likelihood ratio derived through k-fold cross validation was 0.40 for the single-factor configured model, 0.62 for the multi-factor configured model, and 0.66 for the IID configured model when fit to 2019 data.

*Supplementary material S4 reference*

Kristensen, K., Nielsen, A., Berg, C. W., Skaug, H., and Bell, B. M. (2016). TMB: Automatic Differentiation and Laplace Approximation. Journal of Statistical Software, 70(5), 1-21. doi:10.18637/jss.v070.i05

*Supplementary material S5. Results for 2019 3 and 4-age class STMs*

Table S5.1. Average absolute difference in proportion at age class between 3-age class STM predictions and acoustic trawl survey data in International Pacific Fisheries Commission geographic strata

|  |  |  |  |
| --- | --- | --- | --- |
|  | Single-factor STM | Multi-factor STM | IID STM |
| Stratum 0 | 0.45 | 0.45 | 0.44 |
| Stratum 1 | 0.27 | 0.30 | 0.30 |
| Stratum 2 | 0.13 | 0.13 | 0.12 |
| Stratum 3 | 0.09 | 0.04 | 0.03 |
| Stratum 4 | 0.06 | 0.07 | 0.07 |

Table S5.2. Average absolute difference in proportion at age class between 4-age class STM predictions and acoustic trawl survey data in International Pacific Fisheries Commission geographic strata

|  |  |  |  |
| --- | --- | --- | --- |
|  | Single-factor STM | Multi-factor STM | IID STM |
| Stratum 0 | 0.35 | 0.35 | 0.31 |
| Stratum 1 | 0.20 | 0.21 | 0.21 |
| Stratum 2 | 0.08 | 0.07 | 0.06 |
| Stratum 3 | 0.04 | 0.03 | 0.04 |
| Stratum 4 | 0.03 | 0.10 | 0.06 |

*Supplementary material S6. Plots of comparisons between STMs and ATS proportion at age class in 2019*

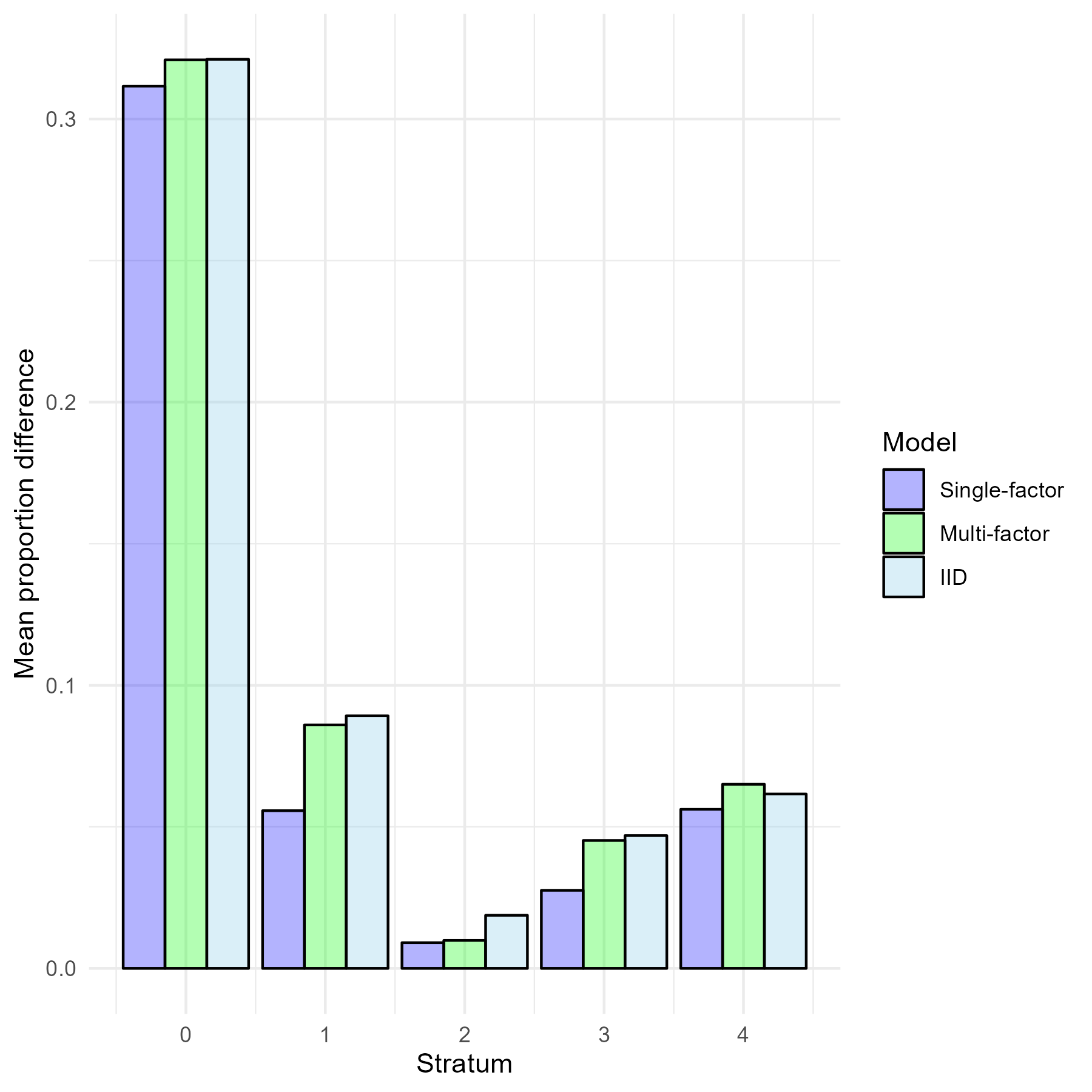


Figure S6.1. Absolute difference in age composition across all age classes (2, 3, 4-6, 7-9, 10+) between different STM configurations and ATS data from 2019.

C:\Users\Derek.Bolser\Documents\Saildrone postdoc\Manuscripts\Fig. S5.2 Panel of proportions at age_revised_multi-factor.tiff

Fig. S6.2. Proportion at age class for the hake survey (ATS) and multi-factor STM in 2019 by INPFC geographic strata. A) Stratum 0, B) Stratum 1, C) Stratum 2, D) Stratum 3, E) Stratum 4

C:\Users\Derek.Bolser\Documents\Saildrone postdoc\Manuscripts\Fig. S5.3 Panel of proportions at age_revised_IID.tiff

Fig. S6.3. Proportion at age class for the hake survey (ATS) and IID STM in 2019 by INPFC geographic strata. A) Stratum 0, B) Stratum 1, C) Stratum 2, D) Stratum 3, E) Stratum 4

*Supplementary material S7. Plot of comparisons between STMs and ATS proportion at age class across years*

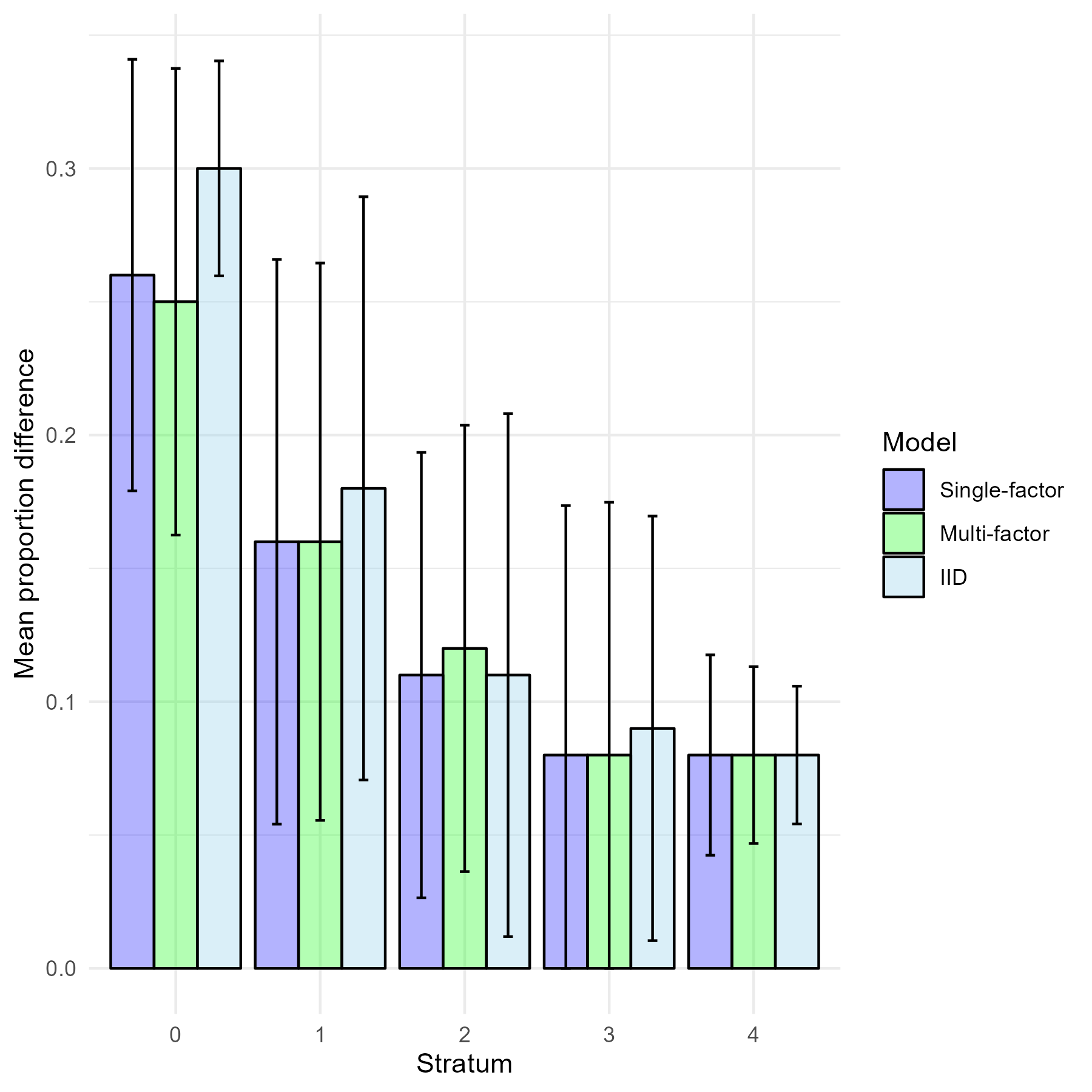


Figure S7.1. Absolute difference in age composition across all age classes (2, 3, 4-6, 7-9, 10+) between STM predictions and ATS data by geographic strata according to model complexity configuration from 2003-2019. Bars show the standard deviation of difference in age composition.

*Supplementary material S8. Tables of average absolute difference in proportion at age for specific years*

Table S8.1. Average absolute difference in proportion at age class between STM predictions and acoustic trawl survey data in International Pacific Fisheries Commission geographic strata in 2003

|  |  |  |  |
| --- | --- | --- | --- |
|  | Single-factor STM | Multi-factor STM | IID STM |
| Stratum 1 | 0.05 | 0.05 | 0.05 |
| Stratum 2 | 0.01 | 0.01 | 0.01 |
| Stratum 3 | 0.02 | 0.02 | 0.01 |
| Stratum 4 | 0.05 | 0.05 | 0.05 |

Table S8.2. Average absolute difference in proportion at age class between STM predictions and acoustic trawl survey data in International Pacific Fisheries Commission geographic strata in 2007. The IID STM did not converge in 2007.

|  |  |  |  |
| --- | --- | --- | --- |
|  | Single-factor STM | Multi-factor STM | IID STM |
| Stratum 1 | 0.27 | 0.27 | -- |
| Stratum 2 | 0.19 | 0.19 | -- |
| Stratum 3 | 0.28 | 0.29 | -- |
| Stratum 4 | 0.07 | 0.07 | -- |

Table S8.3. Average absolute difference in proportion at age class between STM predictions and acoustic trawl survey data in International Pacific Fisheries Commission geographic strata in 2009

|  |  |  |  |
| --- | --- | --- | --- |
|  | Single-factor STM | Multi-factor STM | IID STM |
| Stratum 1 | 0.10 | 0.12 | 0.12 |
| Stratum 2 | 0.12 | 0.13 | 0.12 |
| Stratum 3 | 0.06 | 0.07 | 0.07 |
| Stratum 4 | 0.09 | 0.08 | 0.08 |

Table S8.4. Average absolute difference in proportion at age class between STM predictions and acoustic trawl survey data in International Pacific Fisheries Commission geographic strata in 2011.

|  |  |  |  |
| --- | --- | --- | --- |
|  | Single-factor STM | Multi-factor STM | IID STM |
| Stratum 0 | 0.35 | 0.35 | 0.35 |
| Stratum 1 | 0.26 | 0.27 | 0.28 |
| Stratum 2 | 0.14 | 0.14 | 0.14 |
| Stratum 3 | 0.13 | 0.13 | 0.11 |
| Stratum 4 | 0.13 | 0.12 | 0.12 |

Table S8.5. Average absolute difference in proportion at age class between STM predictions and acoustic trawl survey data in International Pacific Fisheries Commission geographic strata in 2012

|  |  |  |  |
| --- | --- | --- | --- |
|  | Single-factor STM | Multi-factor STM | IID STM |
| Stratum 0 | 0.29 | 0.27 | 0.28 |
| Stratum 1 | 0.31 | 0.31 | 0.32 |
| Stratum 2 | 0.27 | 0.27 | 0.28 |
| Stratum 3 | 0.24 | 0.24 | 0.24 |
| Stratum 4 | 0.10 | 0.11 | 0.09 |

Table S8.6. Average absolute difference in proportion at age class between STM predictions and acoustic trawl survey data in International Pacific Fisheries Commission geographic strata in 2013. The IID STM did not converge in 2013.

|  |  |  |  |
| --- | --- | --- | --- |
|  | Single-factor STM | Multi-factor STM | IID STM |
| Stratum 0 | 0.22 | 0.23 | -- |
| Stratum 1 | 0.08 | 0.08 | -- |
| Stratum 2 | 0.13 | 0.13 | -- |
| Stratum 3 | 0.03 | 0.03 | -- |
| Stratum 4 | 0.12 | 0.12 | -- |

Table S8.7. Average absolute difference in proportion at age class between STM predictions and acoustic trawl survey data in International Pacific Fisheries Commission geographic strata in 2015

|  |  |  |  |
| --- | --- | --- | --- |
|  | Single-factor STM | Multi-factor STM | IID STM |
| Stratum 0 | 0.25 | 0.24 | 0.26 |
| Stratum 1 | 0.23 | 0.22 | 0.22 |
| Stratum 2 | 0.10 | 0.10 | 0.10 |
| Stratum 3 | 0.09 | 0.06 | 0.07 |
| Stratum 4 | 0.10 | 0.08 | 0.06 |

Table S8.8. Average absolute difference in proportion at age class between STM predictions and acoustic trawl survey data in International Pacific Fisheries Commission geographic strata in 2017. The IID STM did not coverage in 2017.

|  |  |  |  |
| --- | --- | --- | --- |
|  | Single-factor STM | Multi-factor STM | IID STM |
| Stratum 0 | 0.12 | 0.10 | -- |
| Stratum 1 | 0.07 | 0.05 | -- |
| Stratum 2 | 0.06 | 0.06 | -- |
| Stratum 3 | 0.11 | 0.11 | -- |
| Stratum 4 | 0.01 | 0.02 | -- |

*Supplementary material S9: Hake fishery and ATS selectivity*

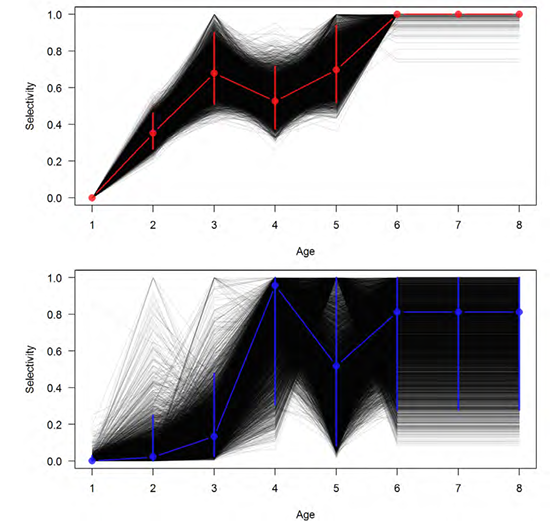


Fig. S9.1. Estimated ATS (top, all years in time series) and fishery (bottom, 2020 only) selectivities from the posterior distribution of the base model. Reproduced from the 2023 hake stock assessment (Berger et al. 2023)

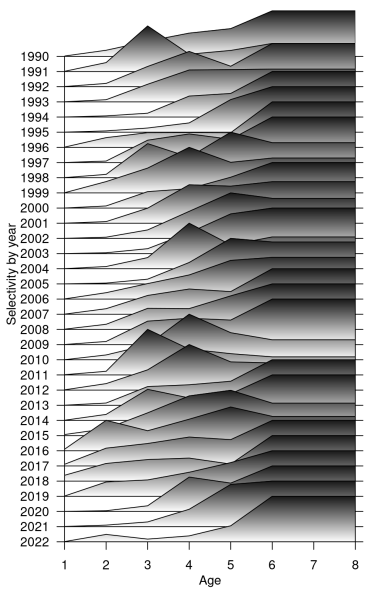


Fig. S9.2. Mountain plot of median fishery selectivity by year. Range of selectivity is scaled to be between 0 and 1 in each year. Reproduced from the 2023 hake stock assessment (Berger et al. 2023)

*Supplementary material S9 reference:*

Berger, A.M., C.J. Grandin, K.F. Johnson, A.M. Edwards. 2023. Status of the Pacific Hake (whiting) stock in U.S. and Canadian waters in 2023. Prepared by the Joint Technical Committee of the U.S. and Canada Pacific Hake/Whiting Agreement, National Marine Fisheries Service and Fisheries and Oceans Canada.

*Supplementary material S10: STM extrapolation grid, spatial knot distribution, and biomass density predictions*

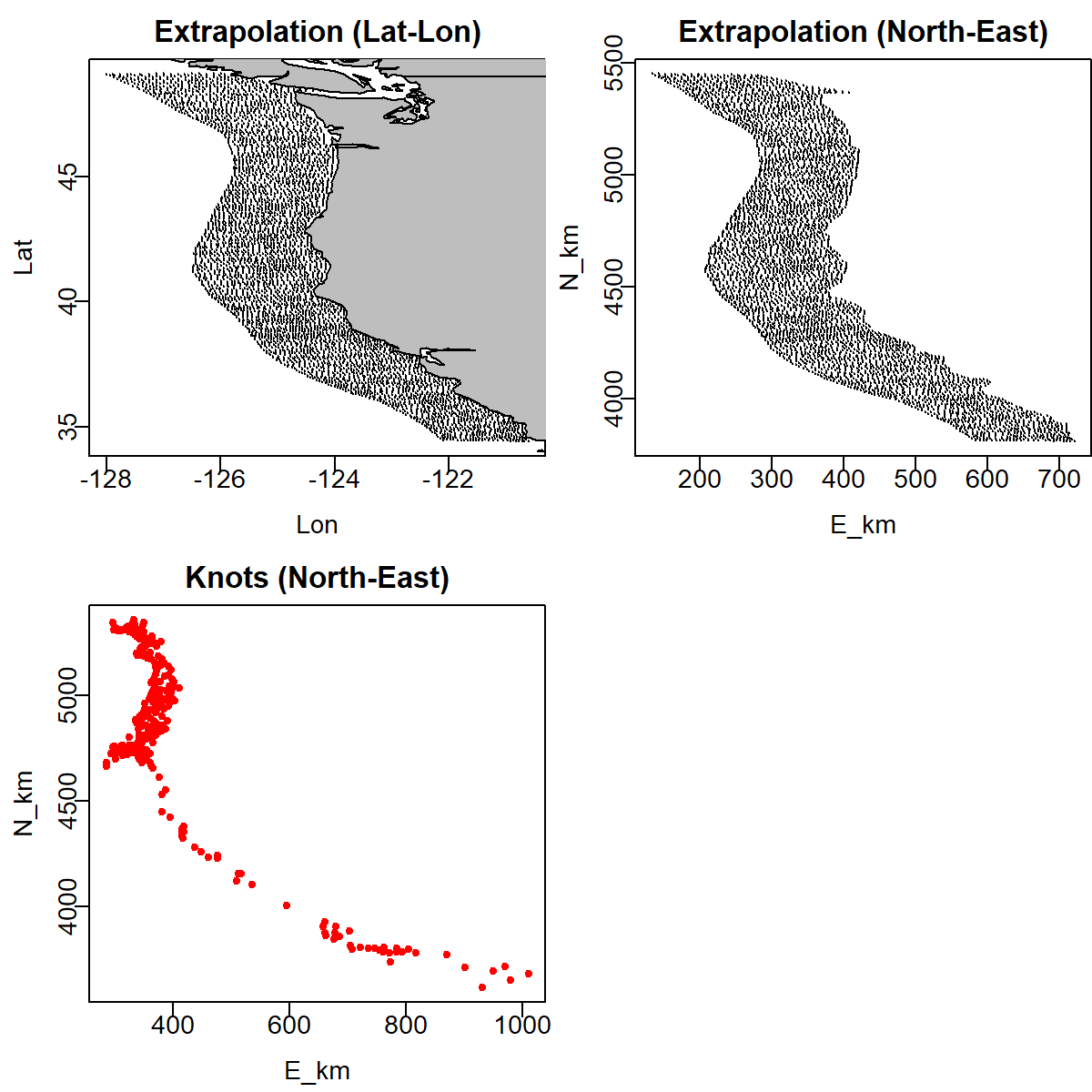


Fig. S10.1. Spatial extrapolation grid and knots employed by STMs

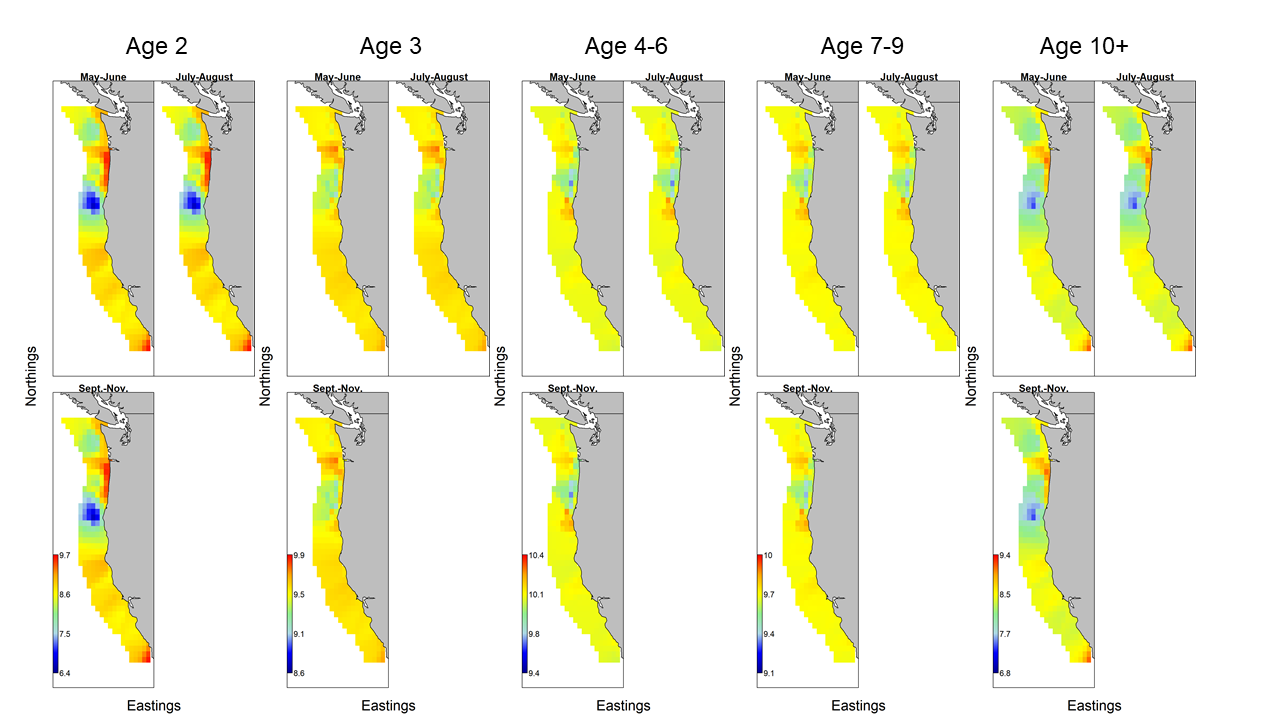


Fig. S10.2. Log biomass density predictions from the single-factor STM

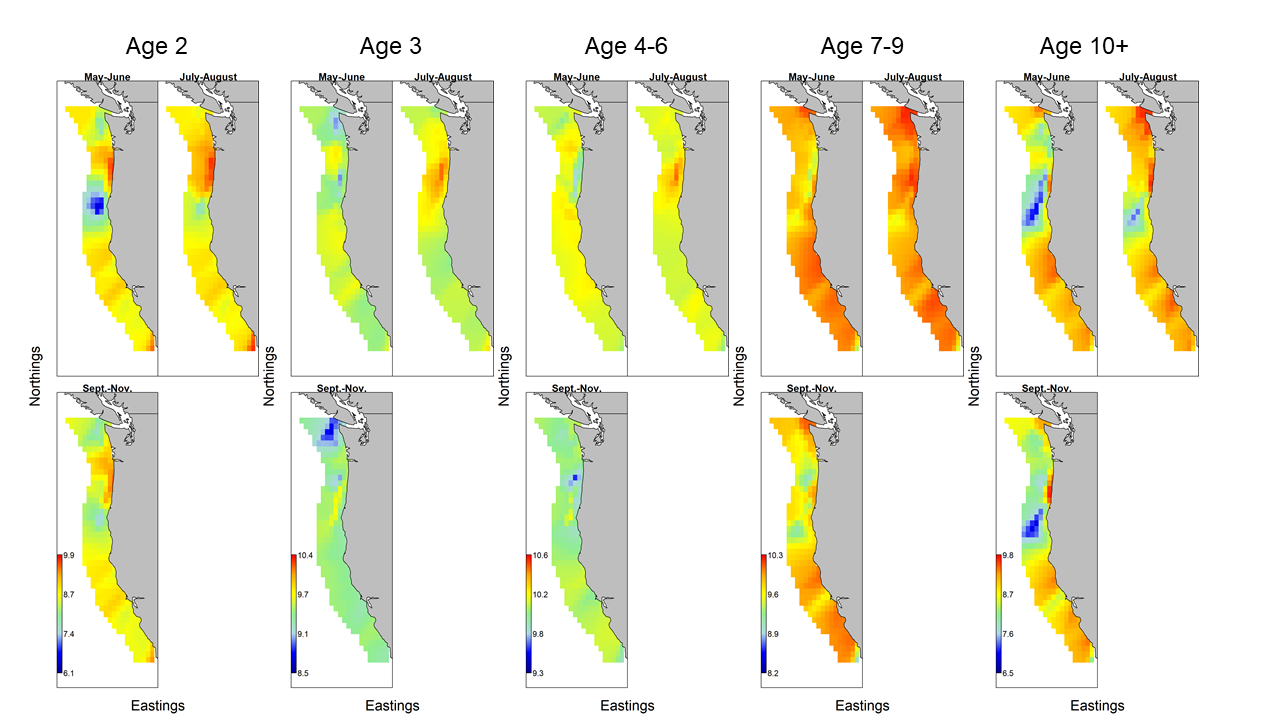


Fig. S10.3. Log biomass density predictions from the multi-factor STM

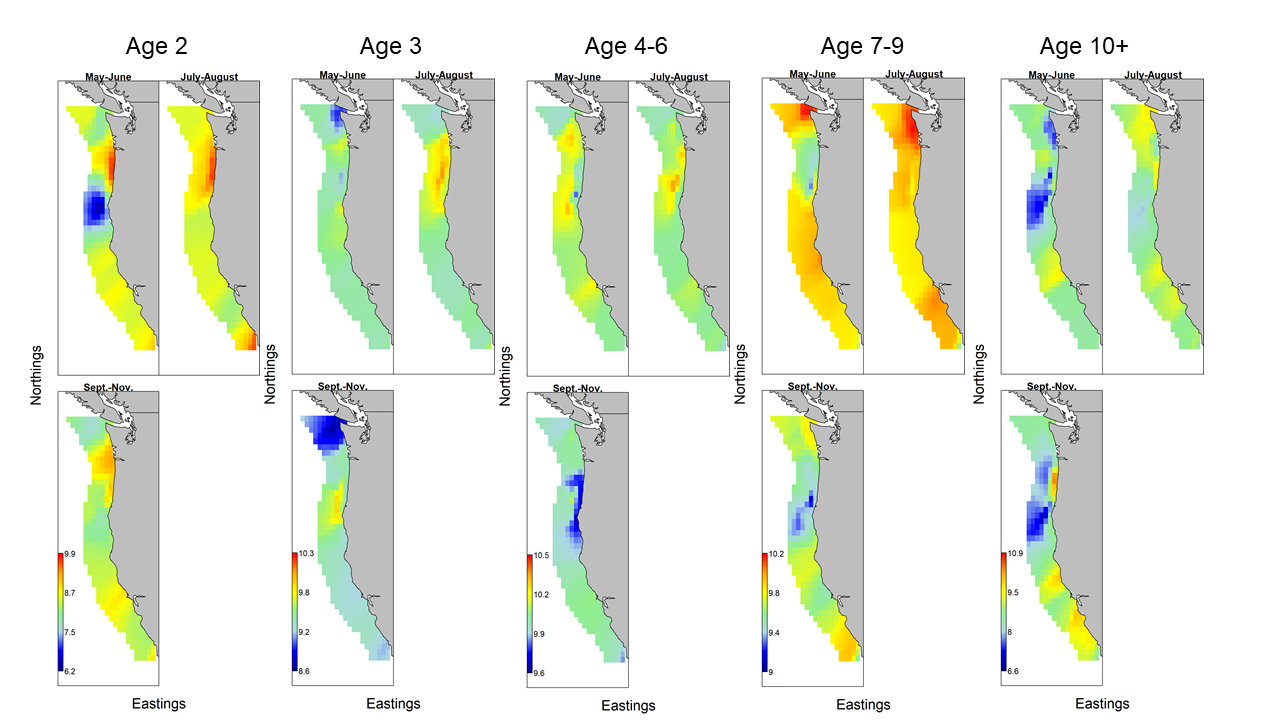


Fig. S10.4. Log biomass density predictions from the IID STM