# Appendix 10c: Summary of the 2021 CIE review of Gulf of Alaska Pacific Ocean perch

The Center for Independent Expert (CIE) review for Gulf of Alaska Pacific ocean perch was conducted virtually from March 30 to April 1, 2021. The panel of experts consisted of Drs Noel Cadigan, Saang-Yoon Hyun, and Geoff Tingley. Overall, the review was productive, resulting in a number of recommendations for future development and research into the assessment for GOA POP. By the conclusion of the review the experts found the assessment to be of high quality, and the reviews contained statements like, “The overall outcome of this assessment, as reviewed, is that it meets the description of best available science and exceeds the acceptability quality threshold to be used to inform management.” (Tingley).

Each of the reviewers provided research recommendations that should serve to improve the assessment model for GOA POP. A number of the recommendations focused on a variety of sensitivity analyses, while others involved more in-depth model development. Distilling these comments, the more in-depth recommendations included:

* Investigate data weighting of compositional data
* Develop a state-space model to be run in parallel to the current assessment
* Continue to investigate use of VAST estimates of survey biomass, in particular investigate reasons behind the divergence between design-based and model-based estimates of abundance

As it pertains to the use of VAST estimates of survey biomass, the consensus among the reviewers was that it is still premature to use this index in the assessment until it can be more thoroughly investigated. This was also the consensus with the use of acoustic survey biomass estimates as an additional index to the model. Due to the recommendations that further work be conducted before implementation into the assessment, and in conjunction with the work that the AFSC internal review team performed through 2020 and 2021 (which additionally identified different methods to estimate fishery selectivity as a topic to be considered in the assessment model development), the GOA POP assessment will not incorporate any substantial model changes for the 2021 assessment cycle, but will investigate and continue to develop these various recommendations to be potentially implemented in the next full assessment that will be conducted in 2023.

Tables 10.21 through 10.25 compile the main recommendations suggested by the reviewers and are organized by the terms of reference (TOR) of the review. A subset of these recommendations were addressed for this Update, and responses to those requests or comments follow the tables.

Table 10.21. TOR 1: Evaluate the data used in the assessments, specifically trawl survey estimates of biomass, and recommend how data should be treated within the assessment model.

| Reviewer | Recommendation | Response |
| --- | --- | --- |
| Tingley | Sensitivities to plausible alternative catch histories, particularly for the early years of the fishery, should be run, but only when there are substantive changes to the assessment model structure or major assumptions. | This sensitivity was investigated for the present cycle (see below). |
| Tingley | Continue to explore different approaches to the appropriate weighting of the composition data, by using different statistical approaches but possibly also by careful quality control of these data, excluding data of known poorer quality. | This has been continually evaluated since 2017, and the results are very sensitive to the biomass index used. We will present updated results in September 2022. |
| Tingley | At a future assessment, it is recommended to try and incorporate all of the high-quality length composition data from both the survey and the commercial fishery, at least in a sensitivity. | We plan to investigate this sensitivity in the summer of 2023 |
| Tingley | Prior to or as part of the next assessment, explore whether the plus group should continue to start at age 25 or whether an older plus group starting age is more appropriate. | We have explored this in previous assessments, but will update this analysis in the summer of 2022. |
|  |  | This sensitivity was investigated for the present cycle (see below). |
| Cadigan | Investigate if stock weights-at-age from the survey are significantly (i.e., in the statistical sense) different than fishery weights-at-age. Also, investigate if there is significant temporal variation in both stock and fishery weights-at-age. Provide figures of how mean weight-at-age changes over time, with different panels for groups of ages (i.e., 1-5, 6-10, 10+). Consider using more efficient and less bias methods for analyzing size-at-age from length-stratified age samples (e.g., Perreault et al., 2019). Investigate spatiotemporal variation in weight as a function of length. | We have previously evaluated time-dependent and have compared between the survey and fishery. We will update this analysis in the spring of 2023, in particular with the different groups of ages, as well as new methods of length-stratified sampling. |
|  |  | This sensitivity was investigated for the present cycle (see below). |
| Cadigan | Consider new sampling programs to collect information on POP maturity. | TBD, dependent on funding |
| Cadigan | Investigate a bootstrap re-sampling procedure (e.g., Jourdain et al., 2020) to estimate uncertainty (i.e., covariance) in survey age compositions. This could also be considered for fishery compositions, although I recognize that it may be less straight-forward if there is data-borrowing for unsampled fishery ?strata? (i.e., gears, areas, seasons, etc.). | Currently being investigated by Siskey et al. results for POP will be presented in September 2022 |
| Hyun | If the survey for the POP stock assessment continues to rely on a bottom trawl survey, they should consider increasing the current trawlable area. | The current method for selecting trawl sites will continue to expand our understanding of trawlable and untrawlable grid cells |
| Hyun | They should revise the calculation of the CV of annual bottom trawl survey indices (annual relative population sizes) because they failed to consider the covariances of survey indices from neighboring strata when calculating the variance of the annual survey index. | We will discuss the potential for this calculation with GAP in the spring of 2022. |

Table 10.22. TOR2: Evaluate the stock assessment model for GOA Pacific ocean perch in general and comment on appropriateness of parameter estimates to assess stock status determinations.

| Reviewer | Recommendation | Response |
| --- | --- | --- |
| Tingley | Exploration of additional information to better define the realistic range of M for Pacific ocean perch is recommended. This should consider data available for Pacific ocean perch and for other long-lived rockfish species. | In the 2020 assessment we used Hamel (2015) as the prior for M. We will be performing sensitivities to M in the summer of 2022, as per the SSC request. |
|  |  | This sensitivity was investigated for the present cycle (see below). |
| Cadigan | Investigate a sensitivity model run with an initial age-structure derived using the assumed M and a few years of F like that estimated for 1961. For example, initial cumulative Z = a*M + min(a,3)*Finit will be appropriate if the stock experienced Finit fishing mortality for three years prior to the start of the assessment model. | Within the internal review team we investigated alternative methods to estimate initial age-structure. We will revisit this with this recommendation in the spring of 2023. |
| Cadigan | Consider including a stock-recruit model with autocorrelated errors to improve the fit of the POP assessment model. Investigate possible drivers of patterns in recruitment deviations. | We have been investigating time-dependent mean recruitment, and will revisit this analysis with this suggestion in the summer of 2022. |
| Cadigan | Consider removing priors for F Regularity and sigma-R. | This sensitivity was investigated for the present cycle (see below). |
| Cadigan, Hyun | A research (i.e., exploratory) state-space stock assessment model, run in tandem with the current stock assessment model, should be developed. | We will begin to develop a state-space model after some of the higher priority suggestions have been addressed. |
| Cadigan | Consider including fishery length composition information in off-years when ages are not measured. However, this may not provide much additional information about recent recruitment trends because of the low selectivity of the fishery for ages less than seven. | We will perform this request as a sensitivity run in the summer of 2023. |
| Cadigan | Evaluate the quality of fishery and survey age compositions for tracking cohorts. | This is a common evaluation in our standard assessments. We feel that given the amount of funding and realistic level of sampling, that our age composition data is adequate to track cohorts. |
| Cadigan | Provide a retrospective analysis of current status evaluations. This will provide additional information on the reliability of the status evaluations. | We will perform this sensitivity analysis in the summer of 2023. |
| Cadigan | Provide convergence diagnostics, including the maximum absolute gradient and the results of a jitter test. | This is potentially a broader topic, but we can fairly easily provide these diagnostics in the 2023 SAFE document. |

Table 10.23. TOR 3: Evaluate the strengths and weaknesses in the stock assessment model for GOA Pacific ocean perch, and recommend any improvements to the assessment model.

| Reviewer | Recommendation | response |
| --- | --- | --- |
| Tingley | In the absence of better information about the likely magnitude of M, sensitivities using values of fixed M that bracket the estimated value M should be run in future stock assessments to inform on the level of risk inherent in the current assumptions about M. | We will perform this sensitivity analysis in the summer of 2022 and present the results of this in September 2022 Plan Team meeting. |
| Hyun | They should incorporate the annual fishery cpue?s into the assessment model framework. | Historically, the fishery CPUE data for POP has been highly variable and questionable, which has caused doubt as to its usefulness in the model. |
| Hyun | They should improve the model fit to the survey indices. One of the efficient ways to improve the goodness-of-fit might be to consider process errors in state variables (random effects). | We intend to develop a state-spaced model once more higher priority model developments are completed. |
| Hyun | The penalized likelihood form as the prior of M, q, and | We will investigate this in the summer of 2023 |
|  | must be revised (beyond the typo). The revised form, which I suggest above, might improve the model performance. |  |
| Hyun | They should do formal model validation, setting true values of free parameters, generating pseudo data, feeding those simulated data into the assessment model, estimating parameters, and comparing estimates of free parameters with the corresponding true values. Such model validation would help us to judge the reliability of parameter estimates and the resultant derived quantities made by the model. | Similar to the model convergence and jitter test diagnostics recommended in the previous TOR, this may be a broader diagnostic to consider in AFSC assessments, however, this model validation will be investigated in the summer of 2023. |
| Hyun | For the retrospective error analysis, they should also examine estimates of annual fishing mortality. | We will perform this sensitivity analysis in the summer of 2023. |

Table 10.24. Evaluate and recommend how survey data are used for biomass indices within the assessment. Specifically, advise on trawl survey indices arising from design-based methods versus model-based approaches.

| Reviewer | Recommendation | Response |
| --- | --- | --- |
| Tingley | Continue to exclude the 1984 and 1987 survey biomass estimates and survey composition data from all future assessments as these are clearly not part of the longer survey timeseries due to the use of differences in vessels, trawl gear, tow duration and survey timing. | We will no longer be including these surveys in the POP assessment. |
| Tingley, Cadigan | Exclude the 1990 and 1993 Gulf of Alaska Bottom Trawl Survey biomass estimates and the survey composition data from all future Pacific ocean perch (and other species) assessments (or include them only in sensitivities, possibly including them as a separate timeseries). These two years do not appear to be part of the longer survey timeseries due to different timing, tow duration and survey structure. | We will investigate the model sensitivity to these surveys in the summer of 2022. |
| Tingley | It is recommended that the current approach of estimating the missing eastern data from the 2001 Gulf of Alaska Bottom Trawl Survey is discontinued for all future assessments of Pacific ocean perch and that one of the proposed approaches, or an alternative approach, is used so as to reduce uncertainty in the next assessment. | We will investigate one of the alternatives in the summer of 2022. |
| Tingley | Continue to support the development and application of spatio-temporal models (such as VAST) for use in stock assessments. In order to make this effective, there need to be a rapid development of a suite of informative diagnostics for spatio-temporal models in a fisheries stock assessment context. Until such time as suitable diagnostics are available, it is recommended that these spatio-temporal models are only used in sensitivity model runs and not in the base case from which management advice is developed. | We will be working with GAP to evaluate the VAST model for POP in the spring of 2022. |
| Cadigan | It was premature to use VAST biomass indices in the POP stock assessment. There are several diagnostic analyses that need to be explored. | We will be working with GAP to evaluate the VAST model for POP in the spring of 2022. |
| Cadigan | Provide the stratum size-weighted averages of the VAST ordinary raw residuals. | We will be working with GAP to evaluate the VAST model for POP in the spring of 2022. |
| Cadigan | Provide trawlable biomass values aggregated over survey strata. This should include time-series of maps indicating strata, where each stratum is colored to indicate the area-expanded VAST biomass. Also useful are time-series plots of VAST biomass aggregated over sets of strata for standard depth ranges shown in Table 2. It will also be informative if this could be further divided into trawlable and untrawlable grounds. | We will be working with GAP to evaluate the VAST model for POP in the spring of 2022. |
| Cadigan | Account for potential vessel and tow time effects in a VAST model. Examine the statistical significance of vessel and tow duration effects. Consider including vessel as a random effect. | We will be working with GAP to evaluate the VAST model for POP in the spring of 2022. |
| Cadigan | Consider including the 1984 and 1987 survey catches in the VAST model, to extend the survey biomass indices back to those years. This VAST model should include those effects that were different or less standardized in the 1984 and 1987 surveys. Consider the potential confounding of year effects with other effects. | We will be working with GAP to evaluate the VAST model for POP in the spring of 2022. |
| Cadigan | Investigate methods to produce length and size compositions that are weighted by VAST spatial density estimates. | We will be working with GAP to evaluate the VAST model for POP in the spring of 2022. |

Table 10.25. Evaluate abundance estimates from summer acoustic-trawl data, and recommend how it may be used within the assessment.

| Reviewer | Recommendation | response |
| --- | --- | --- |
| Tingley | It is recommended that attempts to develop an acoustic abundance index for Pacific ocean perch from the MACE Acoustic Survey data for use in assessments should be discontinued until the evidence base supports a substantially increased likelihood that the processed acoustic backscatter represents a reliable abundance index for Pacific ocean perch. | We will be continuing to work with MACE in 2022 and 2023 to investigate the utility of this survey in the POP assessment. |
| Tingley | It is, however, also recommended that the existing MACE acoustic and trawl data are further explored in detail to ascertain whether the backscatter data can be reliably and robustly be decomposed into Pacific ocean perch and other species or not. | We will be continuing to work with MACE in 2022 and 2023 to investigate the utility of this survey in the POP assessment. |
| Cadigan | More years of acoustic survey data are needed before deciding how it could be included in the POP assessment. However, having an additional fishery-independent abundance index, and in particular an acoustic survey of the off-bottom (i.e., 0.5m) water column, can be quite valuable for detecting changes in availability of POP to the bottom-trawl survey. | We will be continuing to work with MACE in 2022 and 2023 to investigate the utility of this survey in the POP assessment. |
| Cadigan | Continue and improve research on the sources of uncertainty and possibly bias listed above. This should include quantification and incorporation of these sources of uncertainty into acoustic biomass and age/size compositions. | We will be continuing to work with MACE in 2022 and 2023 to investigate the utility of this survey in the POP assessment. |

## Responses to Selected CIE Comments from Spring 2021

## Alternative Catch Histories

Tingley: *“Sensitivities to plausible alternative catch histories, particularly for the early years of the fishery, should be run, but only when there are substantive changes to the assessment model structure or major assumptions.”*

**Response:** Revisiting the historical catch reconstruction would be onerous given that no new historical data sources have emerged since this review, and there is little complementary data (aside from the length compositions) for the early period of the model to corroborate any alternative trajectories.

To address this comment, we leveraged the fact that the base model already separates the data weights assigned to the early (pre-1977) and late (1977-2021) catch time series. In the base model, these series are weighted identically. We explored alternative weights for the early time series of 20%, 50%, and 150% of the value used in the base model, effectively investigating the impacts of reducing or increasing the certainty of this data source. The terminal spawning and total biomass estimates from these models ranged by less than 10%. As expected, the early series and uncertainty thereof affects the model’s estimate of initial and unfished biomass, with less certain (down-weighted) trajectories resulting in slightly higher estimates of these values (Figure 10.20).

There were two additional sensitivities run (results not shown) that provide further insight into this topic. Firstly, a sensitivity where the model begins during the “late” period (1975) – ignoring all data (catches & lengths) from the early catch period – resulted in a population trajectory nearly five times as high as the base model (in terms of total and summary biomass). A separate model run where the early length composition data were dropped, but the historical catches and model start year were the same as the base model, resulted in a perception of unfished biomass that was ~50% higher than the base model, though the population trajectory from ~1975 to present was nearly the same as the base model.

These findings suggest that there is indeed information contained within the early catch series regarding model scale, particularly when contextualized by the early length composition data. Reducing the weight of these data results in qualitatively similar population trajectories, with slightly higher notions of unfished biomass; ignoring these data completely result in a much higher perception of stock size. Given these findings, revisiting the historical catch reconstruction is unlikely to be an influential exercise at this time.

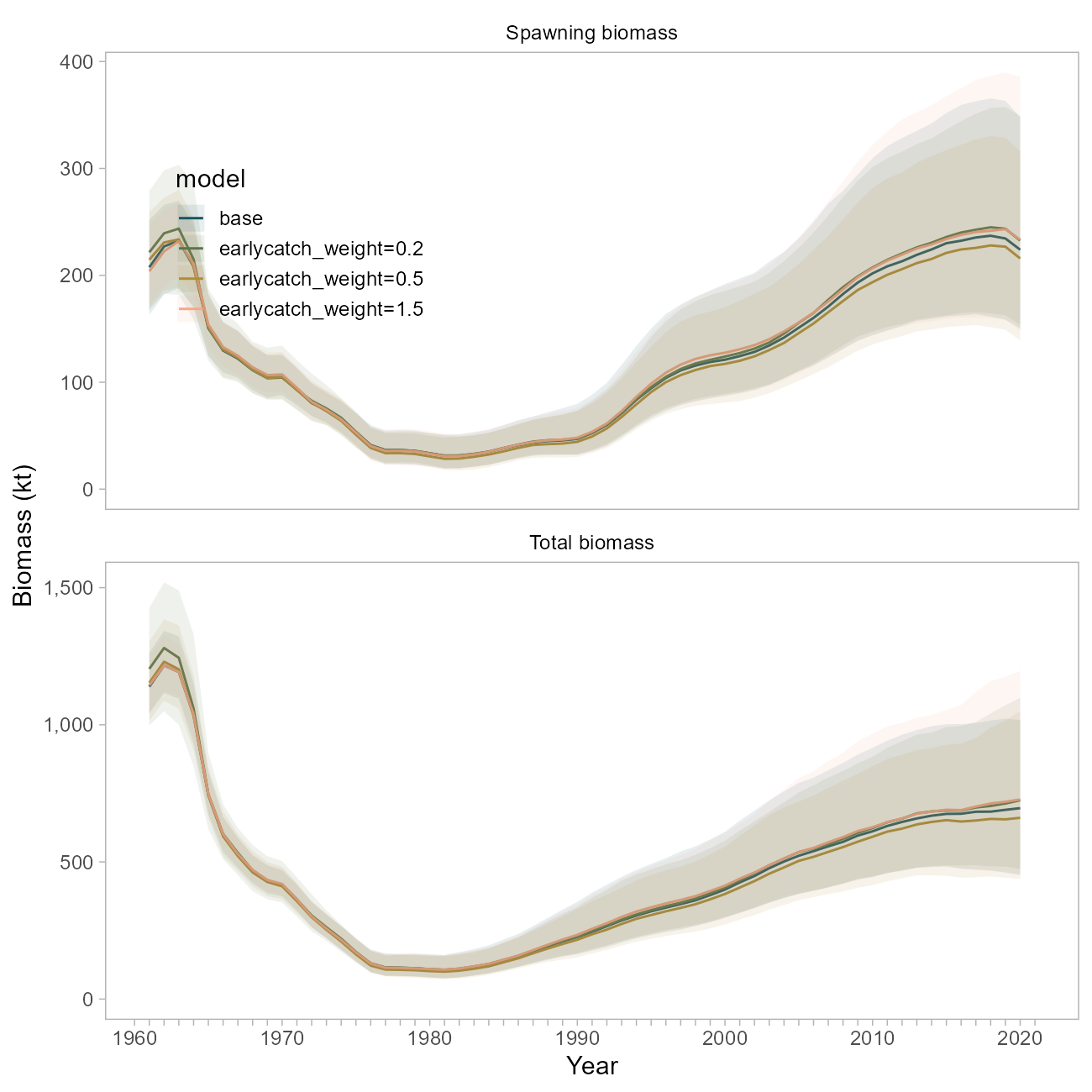


Figure 10.20. Comparison of biomass trajectories between the base model and three sensitivity runs where the early catch time series was weighted at 20, 50 or 150 percent of the weight used in the base model.

## Plus Group

Tingley: *“Prior to or as part of the next assessment, explore whether the plus group should continue to start at age 25 or whether an older plus group starting age is more appropriate.”*

This sensitivity has been explored in previous assessments, and was revisited in a run of the Stock Synthesis version of this model (described below) where the plus group was started at age 29. Model impacts were trivial.

## Stock Weights-at-Age in Survey vs Fishery

Cadigan: *“Investigate if stock weights-at-age from the survey are significantly (i.e., in the statistical sense) different than fishery weights-at-age. Also, investigate if there is significant temporal variation in both stock and fishery weights-at-age. Provide figures of how mean weight-at-age changes over time, with different panels for groups of ages (i.e., 1-5, 6-10, 10+). Consider using more efficient and less bias methods for analyzing size-at-age from length-stratified age samples (e.g., Perreault et al., 2019). Investigate spatiotemporal variation in weight as a function of length.”*

The base model currently uses two size-at-age matrices that represent the probability of a fish of size being age for either an early (pre 1980’s) or late (1980-present) period; both matrices were derived using survey data. Similarly, a single weight-at-age vector developed using survey data is applied to the entire population.

We have previously evaluated time-dependence in size-at-age, and have also previously compared sizes-at-age between the survey and fishery. This analysis is limited by the fact that age records are more sparsely sampled both through time and in terms of overall numbers than the survey (Figure 10.21), which impacts the amount of data available to inform the construction of a separate size-at-age key for the fishery.

To address this comment, we undertook two investigations. First, we ran a model using the original survey-based weight at age vector, but included a new size-at-age matrix for the fishery data only from 1980 onwards. This matrix was defined using the fishery data only, and is more certain as it includes more data from the entire age spectrum (Figure 10.22). This means that fish of all ages, but particularly adults, are more likely to be assigned a length of below 42 cm than the survey-derived matrix would suggest.

The weight-at-age relationship developed using fishery data alone suggests adult fish (ages 20+) to be at a smaller weight in the fishery than in the survey (averaging 742 grams in the fishery vs 891 grams in the survey, Figure 10.23). This would be consistent with discrepancies in selectivity, targeted harvesting, or un-modeled aspects of fisher behavior, *or* could be an artifact of sampling differences between fleets.

Use of this matrix results in slightly lower biomass trajectories (blue line, Figure 10.24), consistent with the notion that the fishery-derived size-at-age matrix assumes a lower probability of larger lengths-at-age.

Separately, we investigated the application of the fishery-derived *weight*-at-age vector (shown in pink in Figure 10.23)). For this sensitivity, the weight-at-age vector was simply replaced with the new values. The biomass trajectory, particularly for spawning biomass, was nearly identical to the base case, much moreso than the sensitivity using the separate size-at-age matrices (green line, Figure 10.24).

This suggests that derived quantities in this model are less sensitive to the weight-at-age parameters than they are to relationship between length and age, and the uncertainty thereof. We believe the survey to be a well-sampled representation of the pouplation, and the associated size-at-age matrix to better represent uncertainty in the growth process for this stock. In this model, this relationship is governed by the size-at-age matrices discussed above, but would be reasonably addressed in a new framework that allows for estimation of von Bertalanffy growth parameters within the model (and the associated variation across ages, or through time).

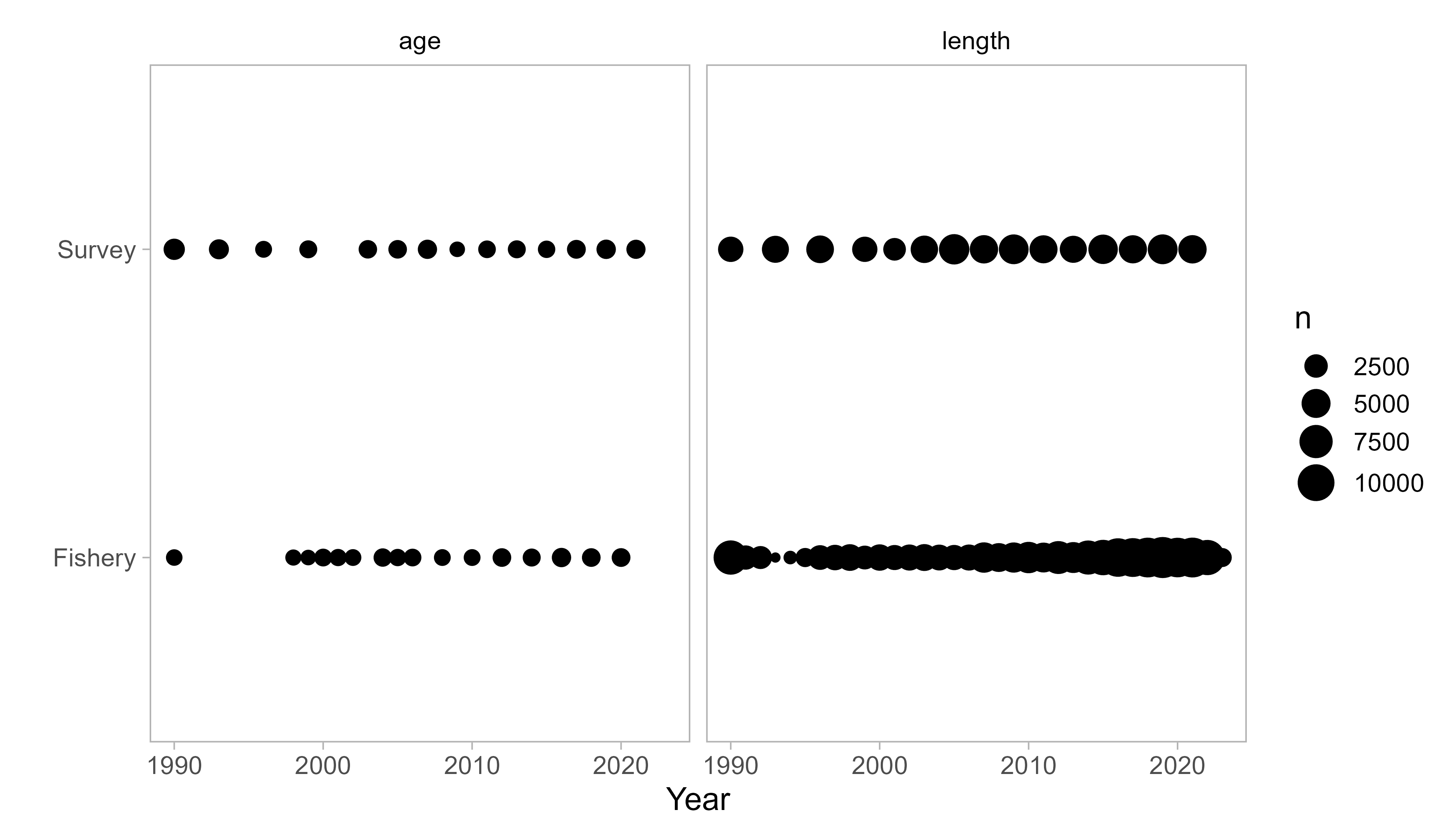


Figure 10.21. Number of raw observations of length and age for the survey and fishery. Note this figure does not represent total data included in the base model, rather the data available for the construction of size-at-age matrices.

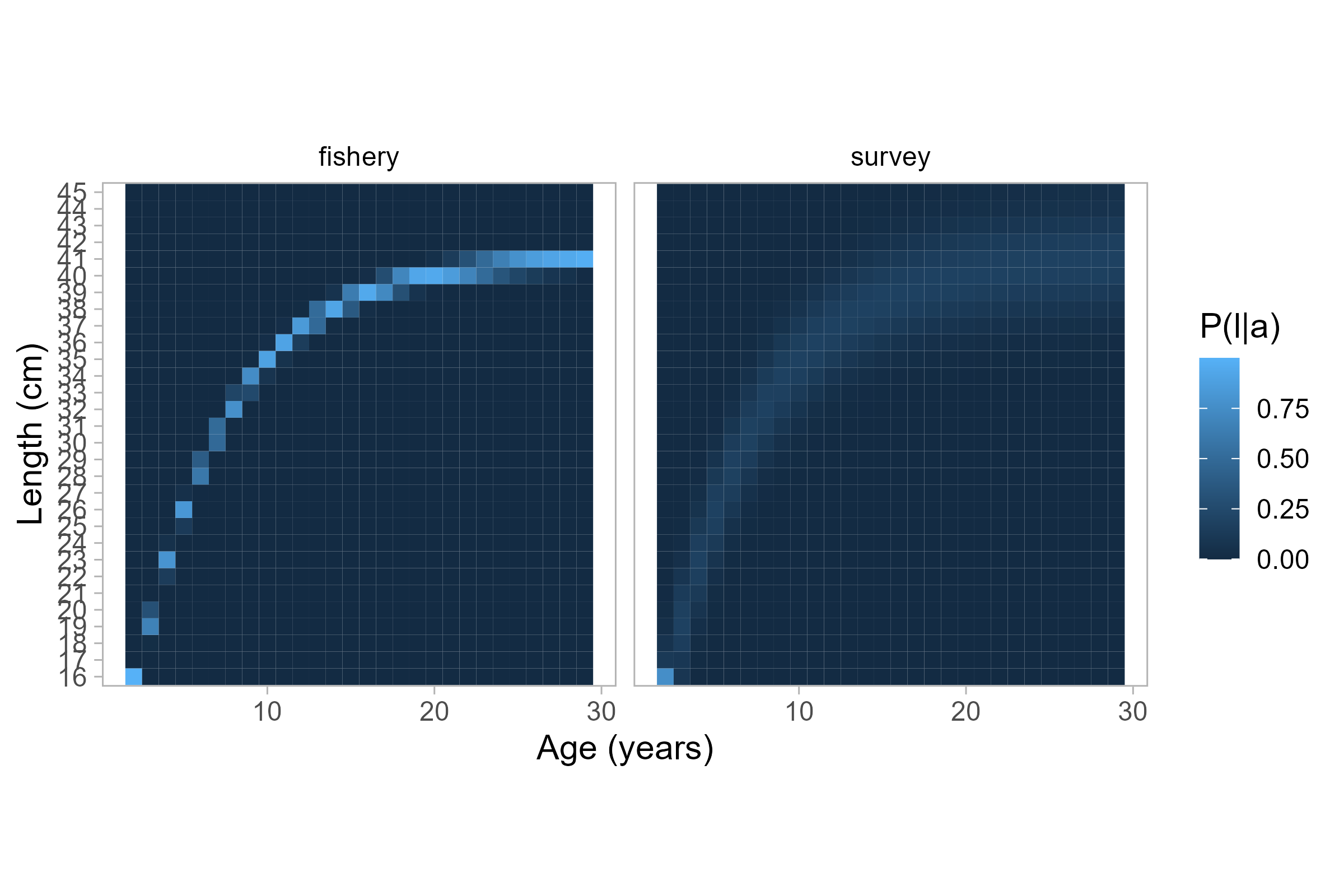


Figure 10.22. Size-at-age probability matrices for each fleet. The matrix on the right is used for all data in the base model.

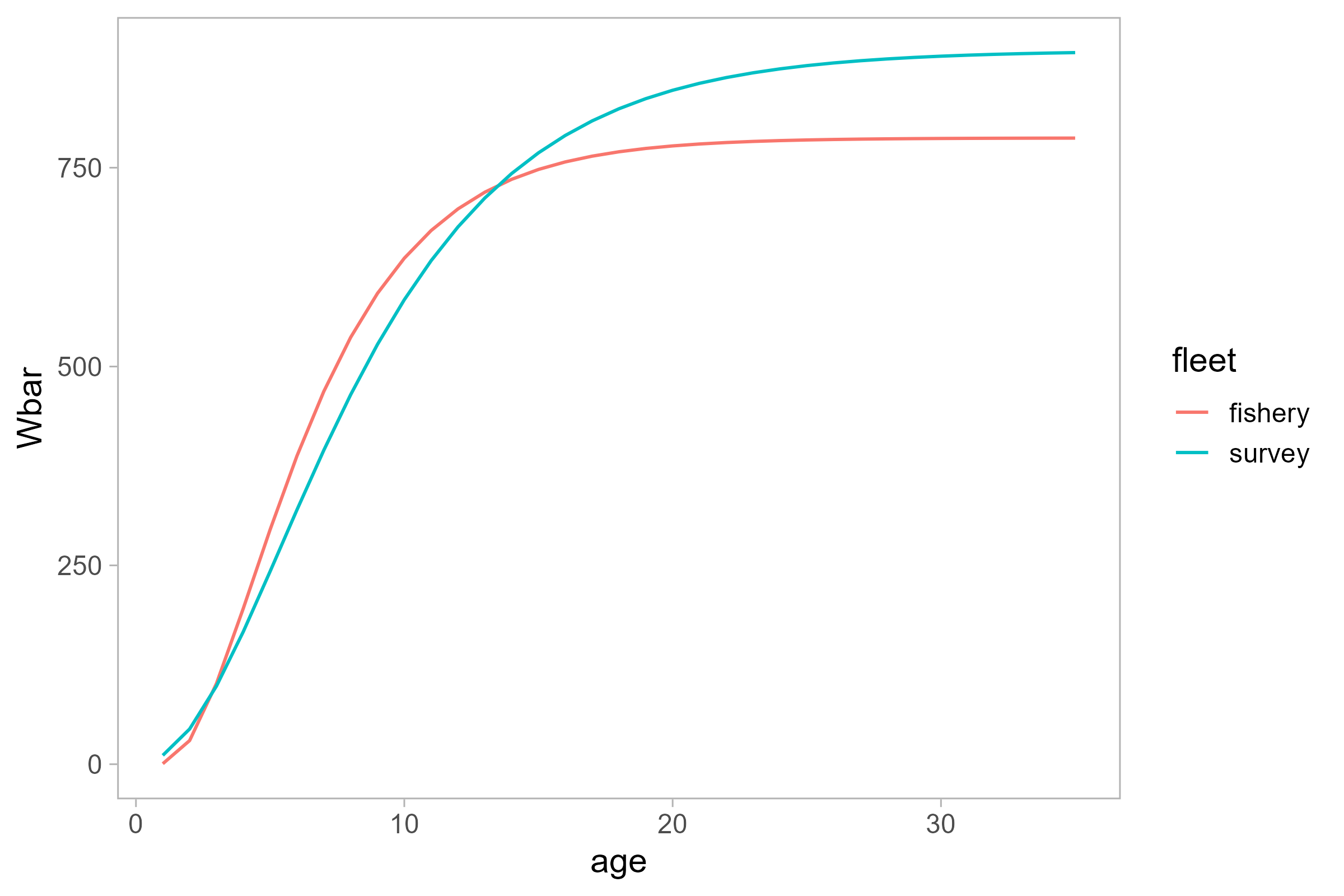


Figure 10.23. Estimated weight-age relationship for two fleets.

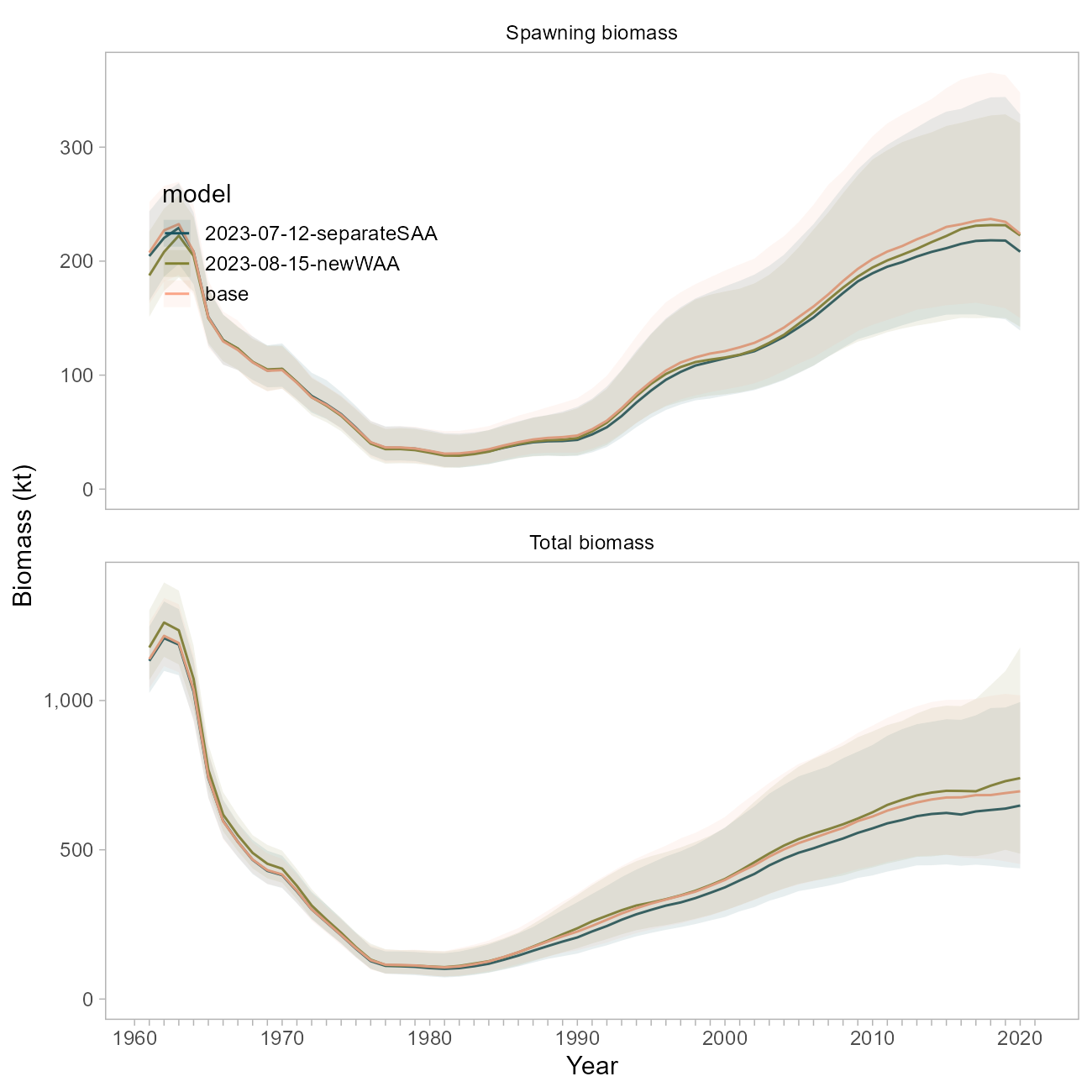


Figure 10.24. Biomass trajectory comparison between the base model, a model using a separate size-at-age matrix for the fishery data from 1980 onwards (‘separateSAA’), and a model using the fishery-derived weight-at-age vector for all population dynamics (‘newWAA’).

## Natural Mortality

Tingley TOR 1: *“Exploration of additional information to better define the realistic range of M for Pacific ocean perch is recommended. This should consider data available for Pacific ocean perch and for other long-lived rockfish species.”*

Tingley TOR 3: *“In the absence of better information about the likely magnitude of M, sensitivities using values of fixed M that bracket the estimated value M should be run in future stock assessments to inform on the level of risk inherent in the current assumptions about M.”*

There has been a fair amount of investigation on this topic; the following response is meant to succinctly describe our findings and is organized into the principal lenses through which was explored in the POP model: through the priors, through likelihood profiles, and through the application of a new modeling framework.

### Priors on M

The current POP base model uses a restrictive maximum-age based Hamel (2015) prior for (a 10% coefficient of variation).

The FishLife R package (Thorson et al., 2023) was recently updated to incorporate morphometric, spawning, behavioral, reproductive and trophic traits from a global database of fish life-history (“FishBase”). This tool enables us to develop an prior for POP that is informed by similar species and therefore better accounts for uncertainty. The prior suggested by the FishLife R package (version 3.0.0) for POP specifically is broader and centered at a higher value (0.0939) base model prior (Figure 10.25). The FishLife prior for the genus *Sebastes* encompasses values from 0 to 0.20. The MLE from the base model (~0.075) is at the ~6th percentile of the POP-specific prior distribution, and the 22nd percentile of the prior distribution for *Sebastes*.

A sensitivity run for the base model using the POP-specific FishLife prior resulted in an even higher estimate of at 0.112 (Figure 10.25). Biomass trajectories from this model (using the broader prior) are higher than the base model (Figure 10.26). The overall NLL from this sensitivity is ~8 units lower than the base model (252.818 vs 260.057 units). The fits to the survey are not visually improved, though the time series of expected survey values is smoother (Figure 10.27).

### Likelihood Profiles on M

We profiled over values of from 0.01 to 0.30 in increments of 0.02 using the base model. This method involves fixing and removing the prior on from the total NLL calculation.

In the absence of the prior, the total likelihood (“objective function”, black line in Figure 10.28) indicates that the MLE for would be much higher than it is in the base model, around 0.15. The data likelihood component (red line in Figure 10.28) is otherwise the most well-defined and suggests an MLE for closer to 0.06, in agreement with the Hamel (2015) prior mean, and consistent with where the base model estimates given the narrow prior and preponderance of information in the data. There was no statistical difference in the data likelihood component for models with values between 0.04 to 0.08.

The fishery size composition data appears to be in conflict with the fishery age composition and survey age and abundance data, whereby the former suggests values much higher than the latter, though both appear to minimize at values outside the tested realm (Figure 10.28).

These observations indicate that the MLE indicated by the data likelihood and the curvature of that profile is probably a compromise between the data sources (survey abundance and ages, fishery ages, and maturity data) that suggest lower, more realistic values for (e.g., less than 0.10) and the one data source (fishery lengths) that suggests a high value for . Data weights are identical for these components, meaning that the higher input sample sizes for the age data are pulling the estimate lower than what the fishery lengths would suggest. The influence of the fishery length data is also reduced in the base model due to the inclusion of multiple other data sources and the specification of the narrow prior.

### Looking at M in a New Modeling Framework

The POP model was transitioned to the Stock Synthesis modeling framework (Methot and Wetzel [2013], v3.30.17). This was *not* undertaken as a full bridging exercise, rather as a learning tool to investigate whether certain issues in the POP model could be resolved or reproduced by changing or simplifying assumptions inherent to the bespoke model framework. This SS model was designed to 1) incorporate all the data that is currently used in the POP base model, 2) better account for uncertainty in key population dynamics processes, with the goal of 3) roughly matching the scale and trend in derived quantities as the base model. *At present, we do not propose the SS model for management use and are not showing extensive model results here*.

Key differences between the 2021 base and SS model include:

* survey catchability is analytical; and
* no size-at-age matrix is used (von Bertalanffy parameters are instead estimated, with attendant uncertainty); and
* *Data Weighting*: we explored a version of the SS model where there are no data weights applied (all data sources’ contributions to the overall objective function are equally weighted, whereas the base model weights all data sources to “1” and the catch data to “50”), as well as a version with three iterations of Francis compositional weights applied. Regardless of whether selectivity was dome-shaped or logistic (see below), the suggested Francis data weights down-weighted all compositional data components. The survey ages were down-weighted the most (between 6% and 9%); the fishery lengths were down-weighted the least (to 16% with logistic selectivity, or 42% with dome-shaped); and the fishery ages were down-weighted in between the two other compositional data sources (to 6% with logistic selectivity and 20% without). The results presented in this appendix use model runs *without* data weights applied. Note that the hessian matrix was invertible for all models, but only upon application of data weights was convergence (maximum gradient < 1e-5) achieved.
* *Fishery Selectivity* is specified differently in the SS model. The functional form is double normal, so that fishery selectivity can be dome shaped , logistic, or somewhere in between (as it is in the base model), given the same four time-blocks as in the base model. We explored either forcing fishery selectivity to be logistic, or allowing the descending limb to form a dome. Dome shaped selectivity resulted in a large increase in model scale (Figure 10.30).

We investigated the influence of in the SS model, using either the original, narrow prior, the broader POP-specific FishLife prior described above, or fixing to the 2021 MLE (0.075, Figure 10.25). These three alternatives were tested against the two selectivity specifications mentioned above.

Several useful findings emerged from this effort:

* The prior seems less influential on model dynamics than does the specification of fishery selectivity, and whether or not data weights are applied. For example, a model using the FishLife prior, logistic fishery selectivity and no data weighting estimated to be 0.032 (Figure 10.29a); when data weighting was enabled for this same model, the estimate agreed more with the prior with an MLE of 0.091 (Figure 10.29b). In an SS model where selectivity was allowed to be dome-shaped, was closer to the 2021 base value at 0.086 (Figure 10.29c)) though the model scale is greatly increased (Figure 10.30).Using the original, restrictive Hamel (2015) prior in a model with logistic fishery selectivity and no data weights resulted in an estimate nearly equivalent to the prior mean (0.061, Figure 10.29d).
* Biomass trajectories across our six experimental runs illustrate that *variation in model scale is most readily described by differences in selectivity* versus the prior or estimate of . Specifically, all models with dome-shaped selectivity exhibited higher biomass trajectories, better fits to the compositional data, and worser fits to the survey data (grey lines, Figures 10.30 and 10.31)), while those with logistic selectivity were closer in scale to the 2021 model, did not fit the compositional data as well, but fit the survey best overall (blue lines, Figures 10.?? and .31)). The tradeoffs in model fit are consistent with observations from the base model that there are conflicts between these data types.
* Recall that using the FishLife prior in the base model resulted in an estimate of to be high at 0.11 (Figure 10.25), which was illustrated further using likelihood profiles on the base model (Figure 10.28). The fact that this dynamic ( pushing ever-higher) did not persist in the SS model is likely *due to the increased flexibility of the double-normal curve*; the base model is controlled by the gamma function to be traditionally dome-shaped. This would explain the need to constrain in the base model when the transition to gamma-shaped selectivity was made; the SS model assumes that older fish are indeed selected by the fishery, and there is information in the fishery and survey ages to suggest that is low.
* We ran likelihood profiles on on the experimental SS model(s) (Figure 10.32). The *conflict between survey and age versus length (fishery) data persists*; the recruitment trend would suggest a higher value of , and this was true for models with logistic or dome-shaped selectivity. This indicates that the conflict between these data sources is not an artefact of the bespoke model. The main distinction between the two profiles is that the model with logistic selectivity effectively ignores the prior and estimates a very low value, likely to compensate for the high exploitation of older fish. The reader is advised that *there was a large range of values that were statistically indistinguishable* for both models (between 0.02 and 0.12).

### Conclusions regarding M

Overall, given the framework of the base model, it is apparent that the current prior on is reinforcing the perception that is between 0.04 and 0.07, and that in the absence of this prior (or in the presence of the broader FishLife prior) the base model’s estimate for would be higher. This estimate represents a compromise between the survey index, survey age, and fishery age data (which all suggest low values for ) and the fishery length data, which suggests a high value for . Using a higher and broader prior in the base model resulted in improved likelihood scores, but presented changes in the scale of the population.

Investigations using an alternative modeling framework revealed that the specification of fishery selectivity seems more influential than in model scale and overall fits. We confirmed via likelihood profile that 1) there is conflict between the compositional data sources and 2) there is a broad range of values (from 0.02 to 0.12) that are statistically indistinguishable. The observation regarding a range of values was also found in assessments of POP for the US West Coast (Wetzel et al., 2013).

The data conflict is much more pronounced in the current POP model versus the SS model, both in terms of the range of values that are statistically indistinguishable, the discrepancy between MLEs for for each likelihood component, and the influence of the prior on the overall MLE. A visual inspection of the data reveals that the fishery length compositions appear nearly stable through time, while the age compositions from both sources appear to track cohorts to a greater degree. Secondary to the recommendations below, we will explore data weighting approaches in the new modeling framework. We re-iterate that the magnitude of disagreement between data sources appears less pronounced in the SS model.

*We do not recommend transitioning to the FishLife prior within the current 2021 model framework*. Though the prior on its own is likely a better represents the uncertainty in POP life history, this decision decision should not be made without concurrently revisiting the treatment of *selectivity, and potential re-weighting of compositional data*. Given our findings above, *we recommend this be undertaken within the context of a new modeling framework*, where it has been revealed that the prior is less influential than other factors.

The previous assessment author indicates that the strict prior on in the base model was a necessary compromise when the fishery selectivity was transitioned to the current gamma (dome-shaped) distribution, so it is likely that both selectivity specification and differences in how population dynamics are represented (e.g., recruitment) between the current and SS models are facilitating estimation of . We conclude that it is worthwhile to continue developing the SS model with particular attention on these processes. *The use of the FishLife prior (or a hybrid of the Hamel and FishLife approaches) is likely appropriate, and not unduly influential, within the SS framework*.

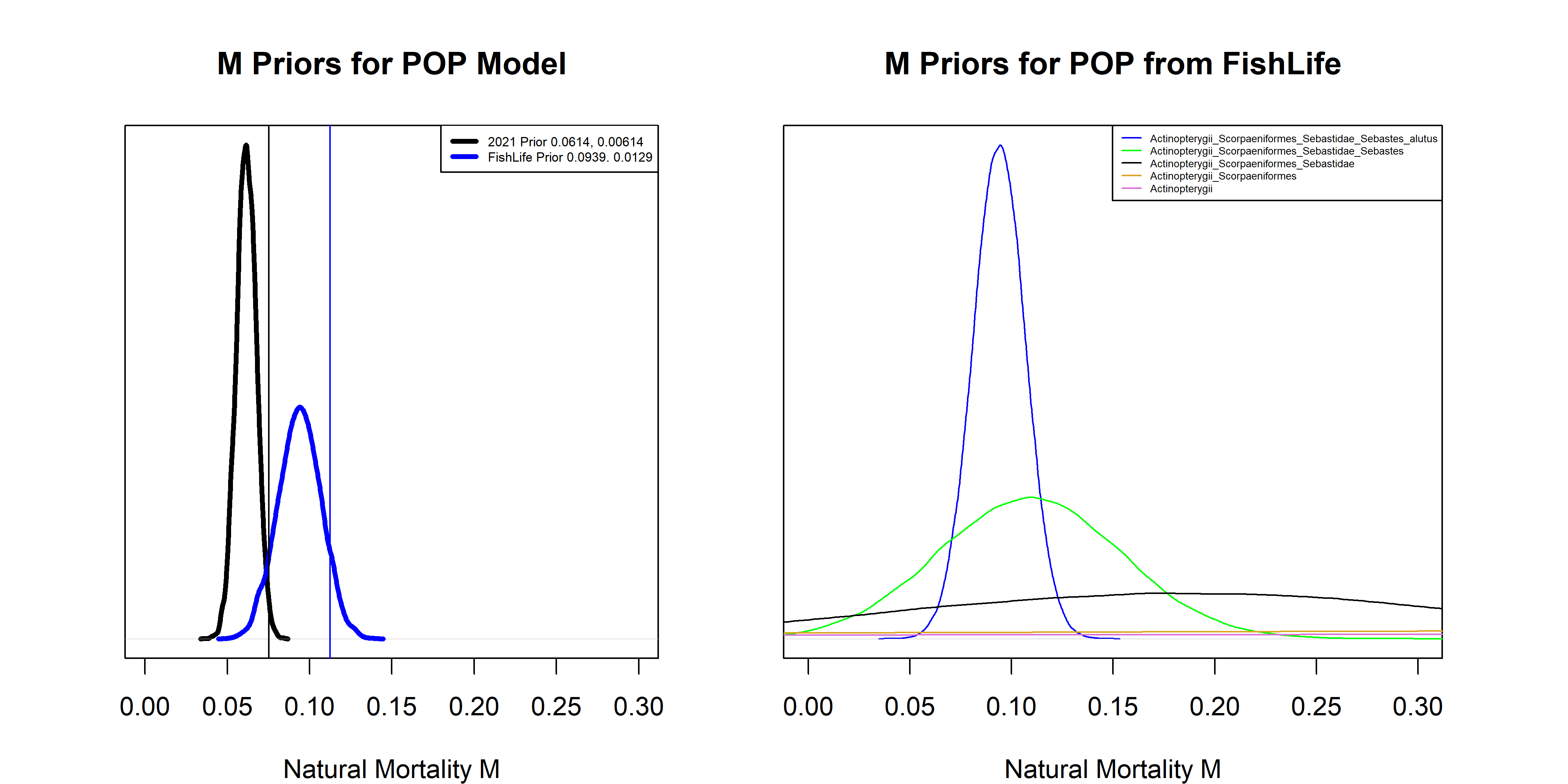


Figure 10.25. Left: Comparison of M priors (thick lines) and maximum likelihood estimates (thin vertical lines) between the base model (black), and the base model using a prior from the FishLife package (blue). Right: M priors from the FishLife package for POP and related taxa.

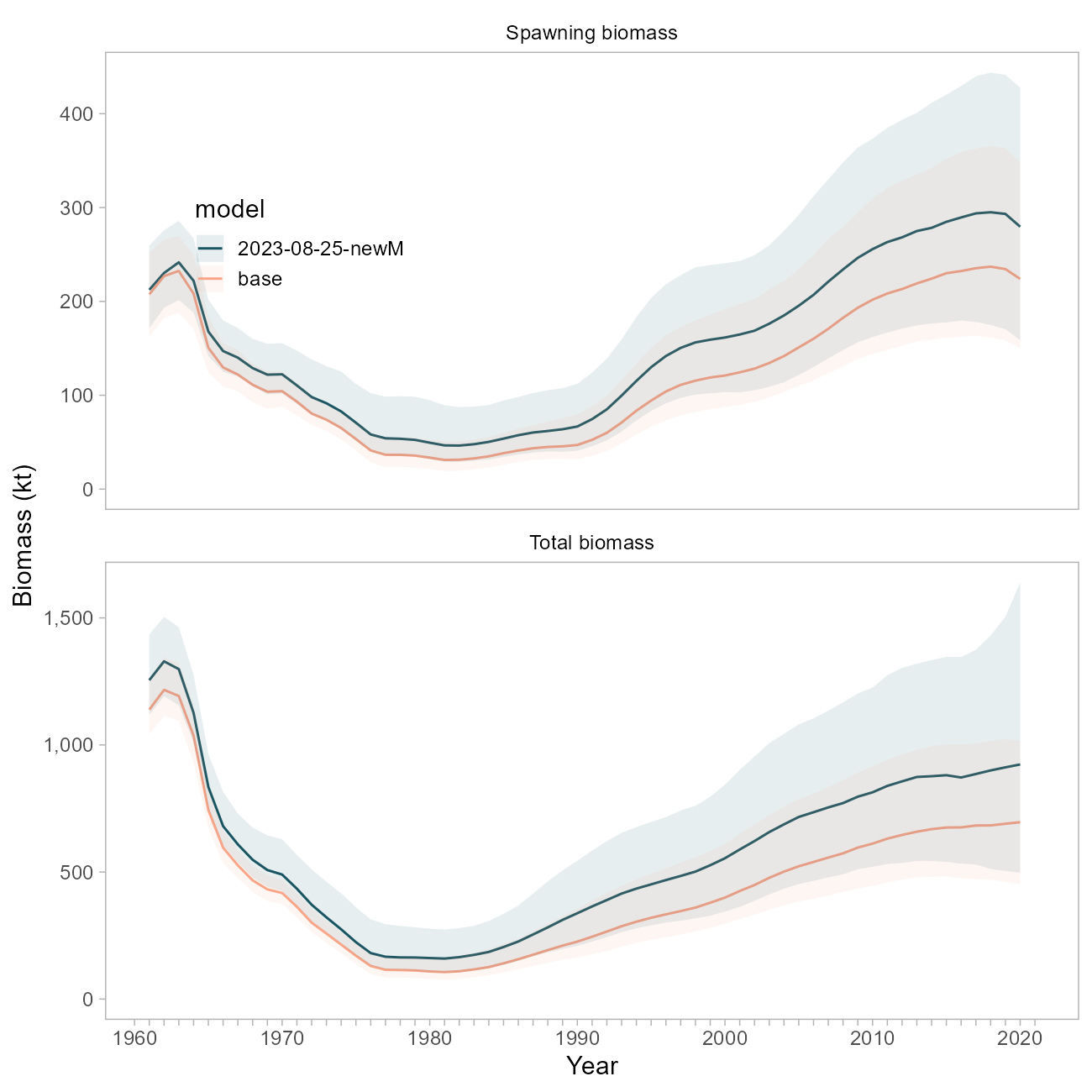


Figure 10.26. Comparison of biomass trajectories between the base model and a model using the FishLife prior for natural mortality.

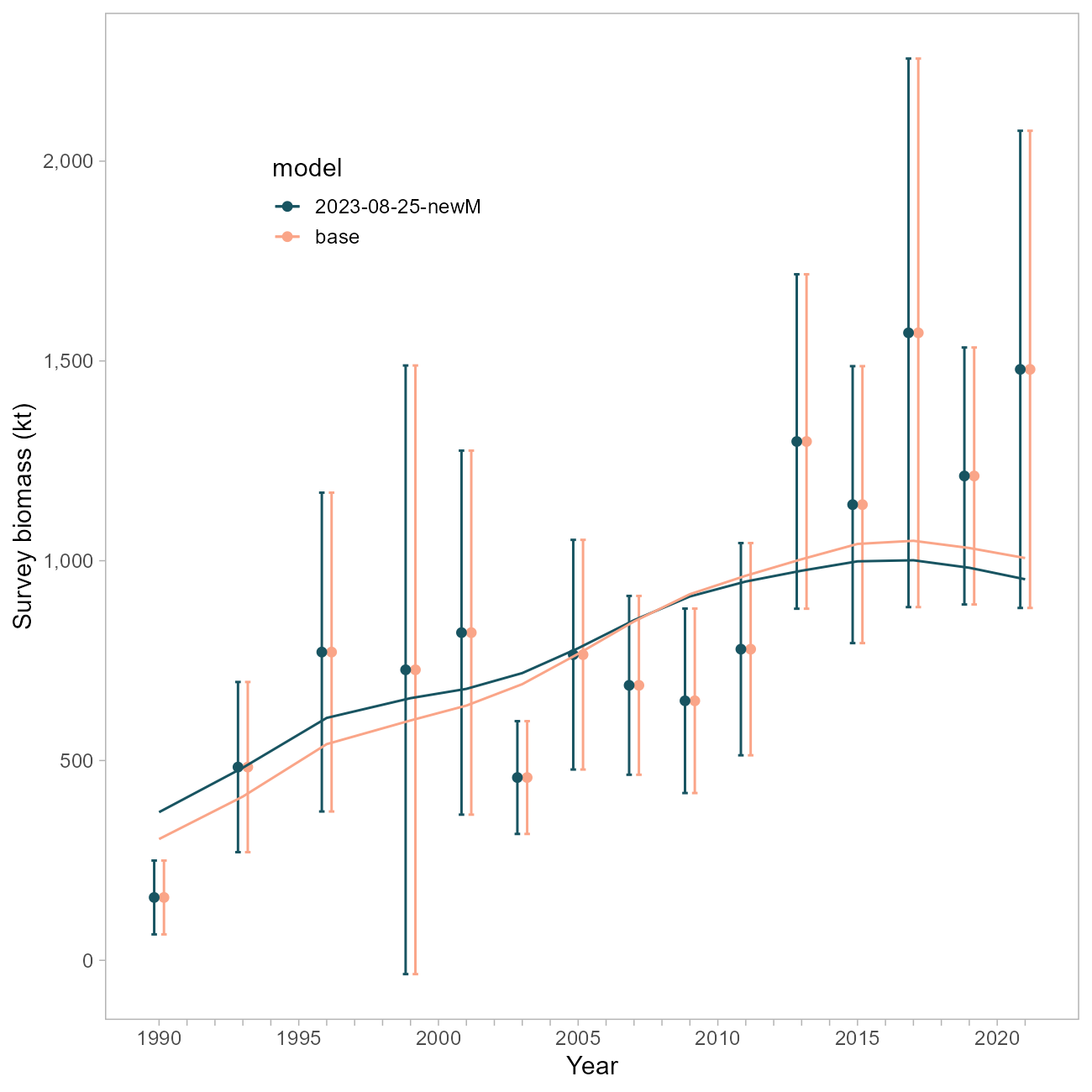


Figure 10.27. Comparison of survey fits between the base model and a model using the FishLife prior for natural mortality.

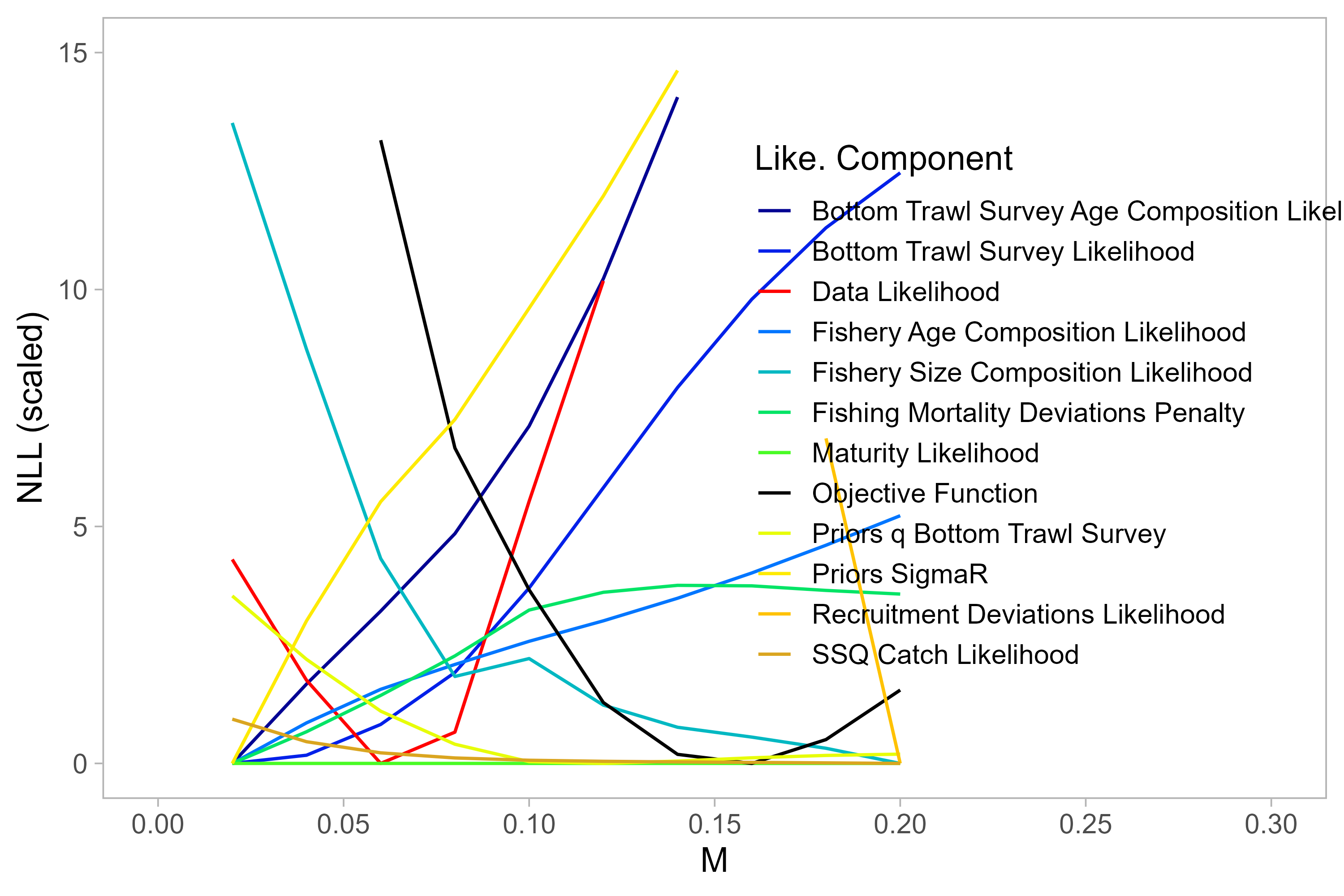


Figure 10.28. Likelihood profile on M using the base model.

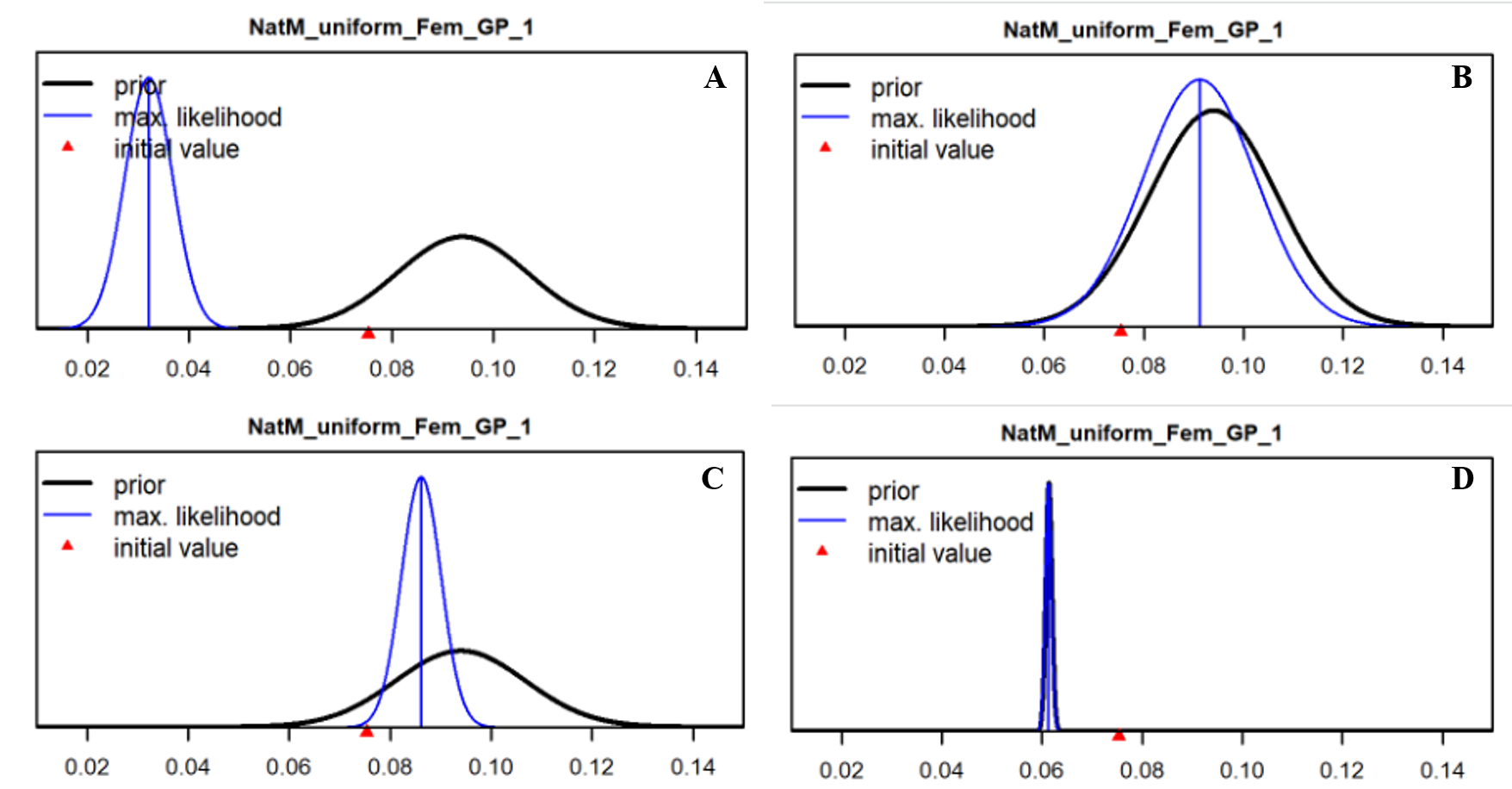


Figure 10.29. Prior (black line) and posterior (blue line) estimates of M using (A-C) the new FishLife prior or D) the original Hamel (2015) prior. Fishery selectivity is forced to be logistic in A, B and D. Tuned compositional data weights using the Francis method have been applied in B.

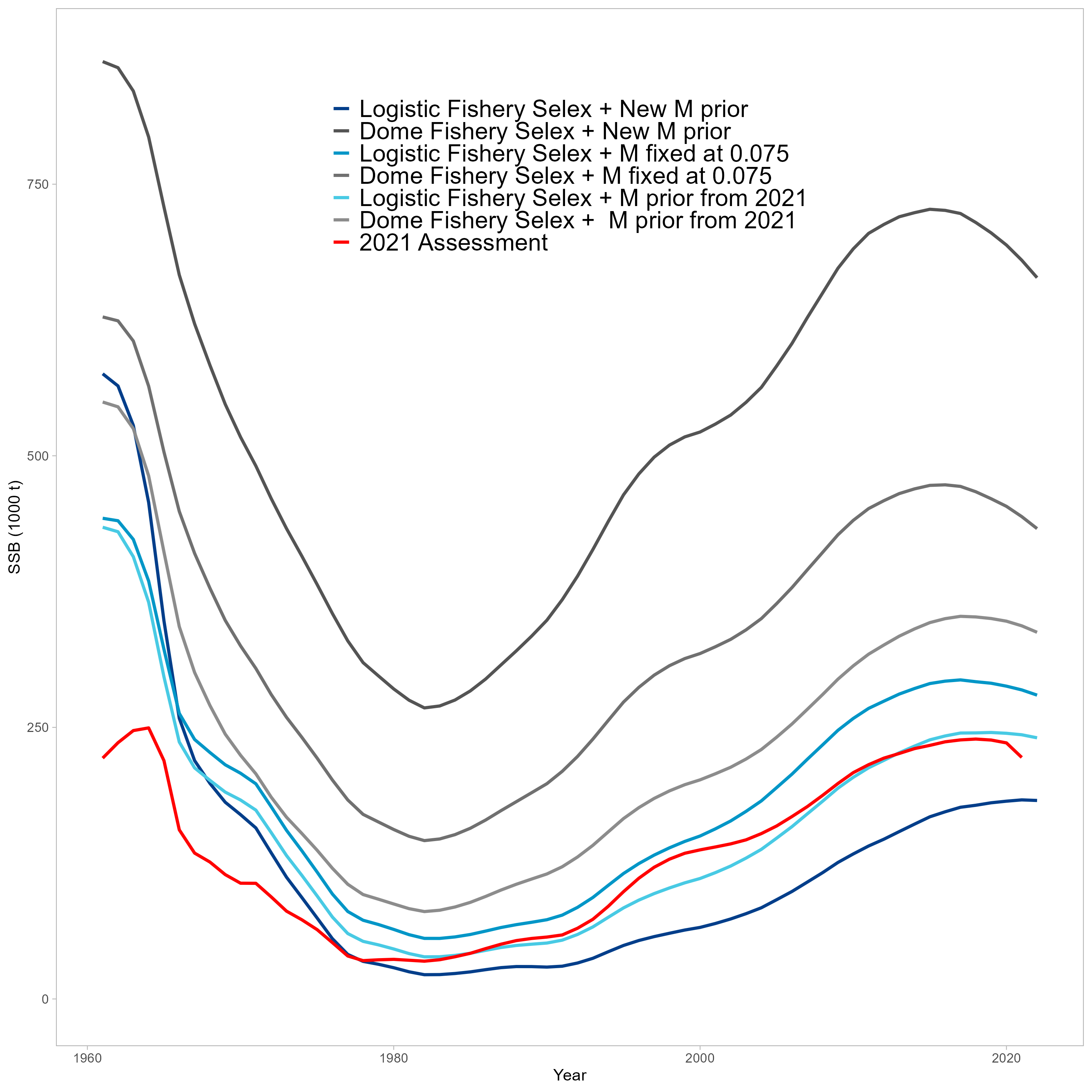


Figure 10.30. Comparison of SSB trajectories for alternative SS configurations implemented for natural mortality explorations. The red line is the 2021 POP Assessment.

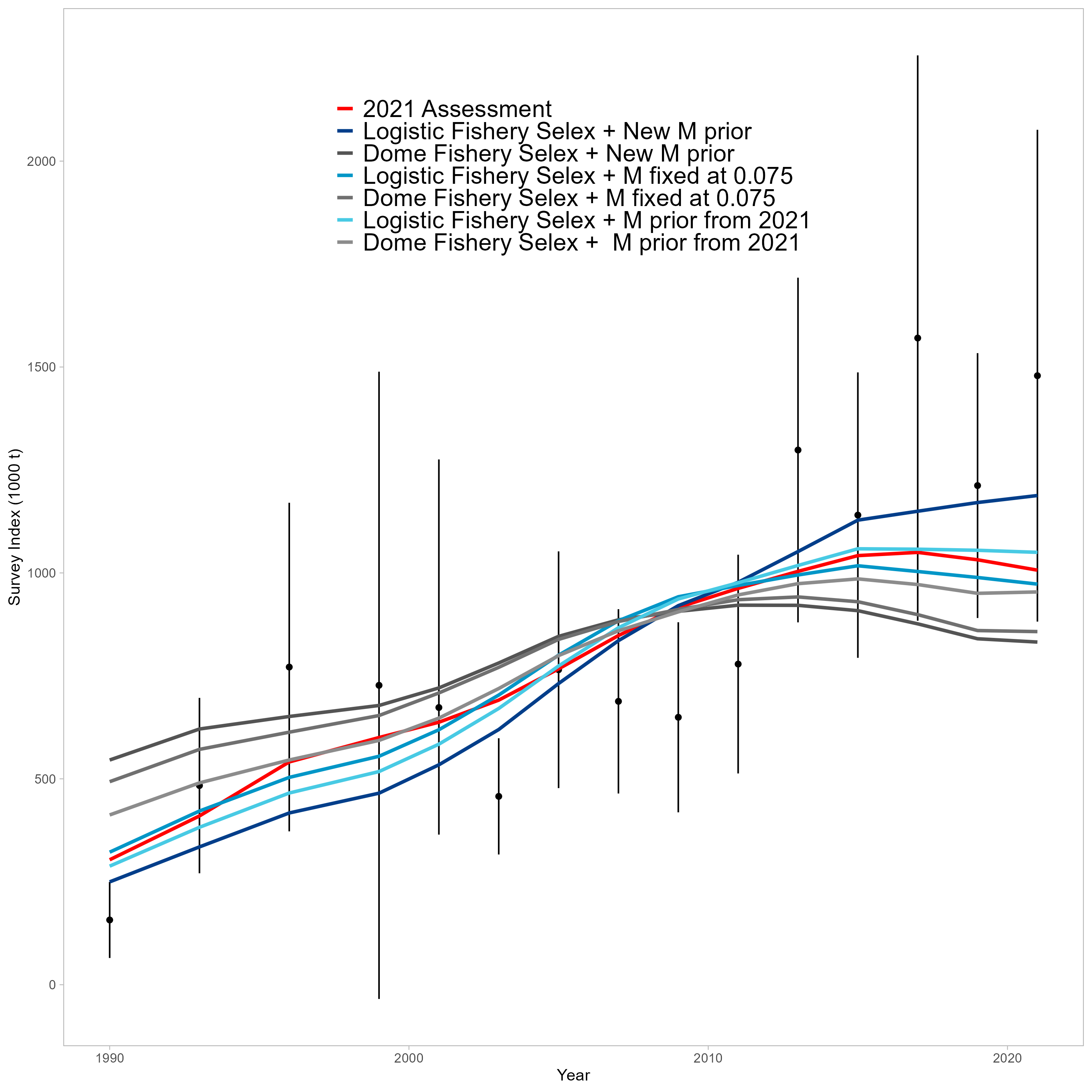


Figure 10.31. Comparison of survey fits for alternative SS configurations implemented for natural mortality explorations. The red line is the 2021 POP Assessment.

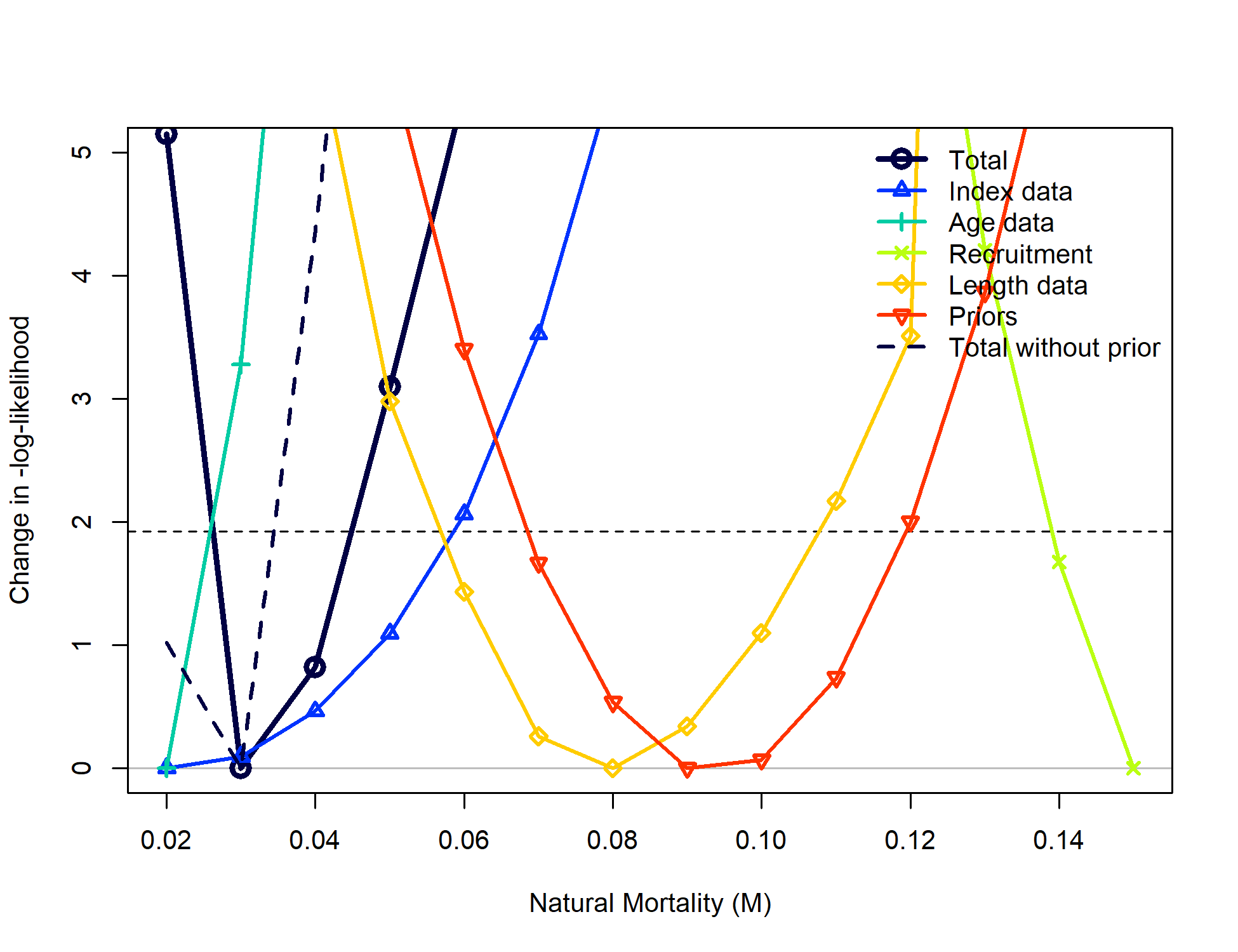


Figure 10.32. Likelihood profile on M using the SS Model. Values below the horizontal dashed line are statistically indistinguishable.

## Priors & Penalties on *F*,

Cadigan: *“Consider removing priors for Regularity and .”*

The “prior for Regularity” a term that penalizes the vector of deviations using the sum-of-squares (in practice, assuming a mean of 0 and a variance of 1). is indeed estimated using a lognormally-distributed prior( ). We addressed this comment by separately disabling each of these functions; additionally, the SS model mentioned above does not involve a prior on nor a penalty for , so comparisons can be made among model frameworks for further information.

Disabling the penalty on did not result in changes to biomass trajectories (Figure 10.33).

Removing the penalty on did result in changes to the biomass trajectories, such that the sensitivity run estimated to be lower and the overall biomass to be higher in the absence of a penalty (Figure 10.34). The SS model, by comparison, estimates to be higher than both ADMB models yet the trajectory is similar (Figure 10.35).

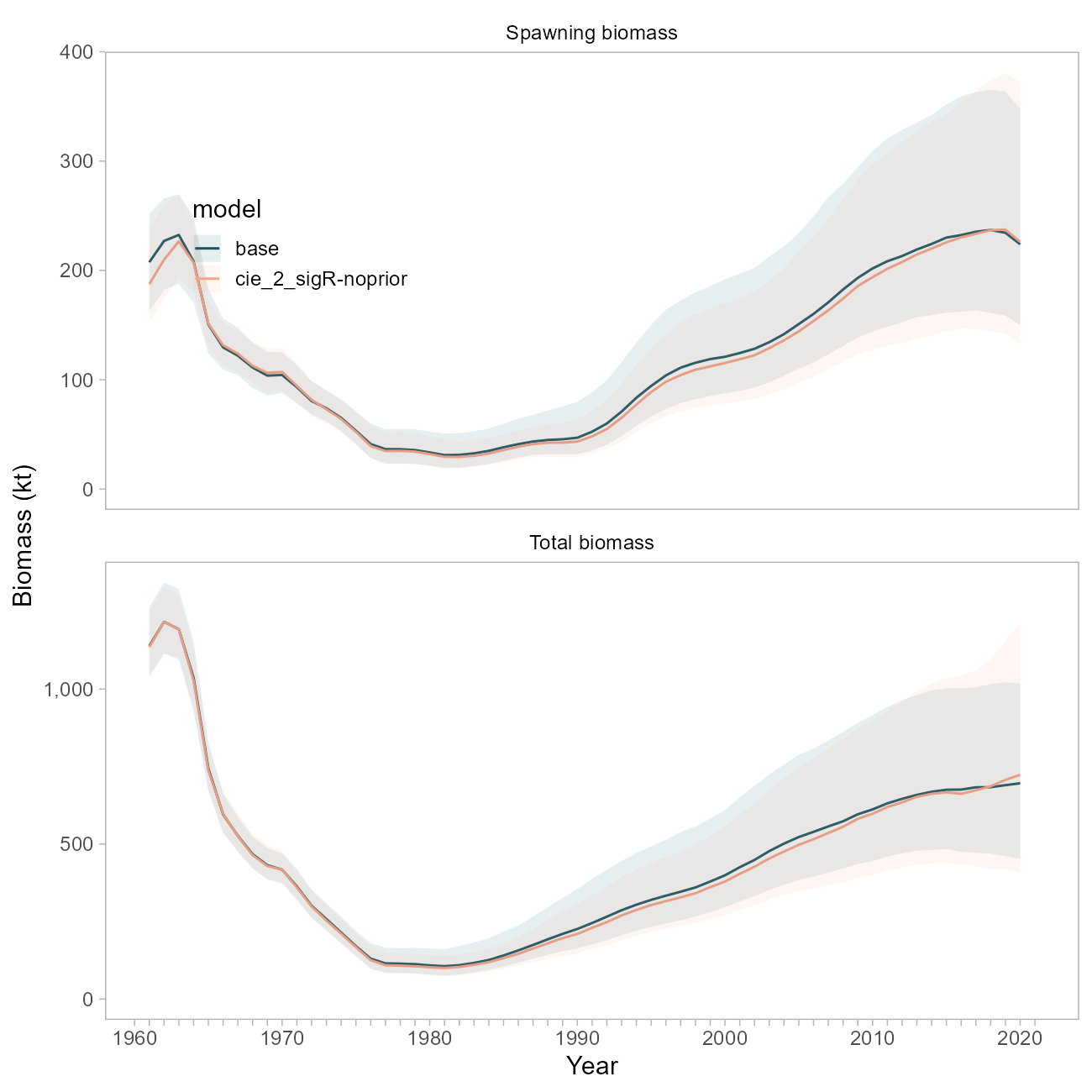


Figure 10.33. Comparison of biomass trajectories between the base model and a model with the prior on sigma-R disabled.

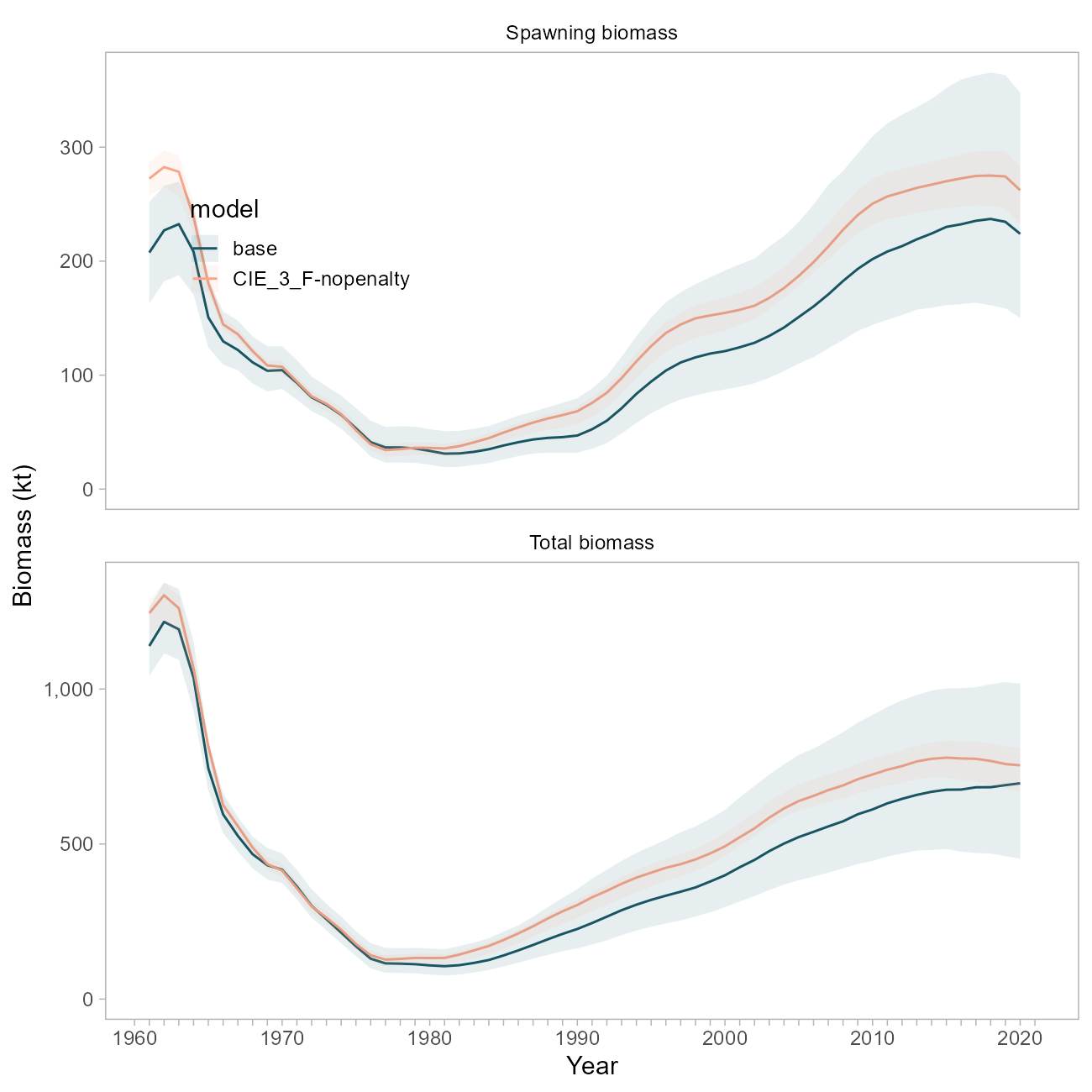


Figure 10.34. Comparison of biomass trajectories between the base model and a model with the regularization penalty on F disabled.

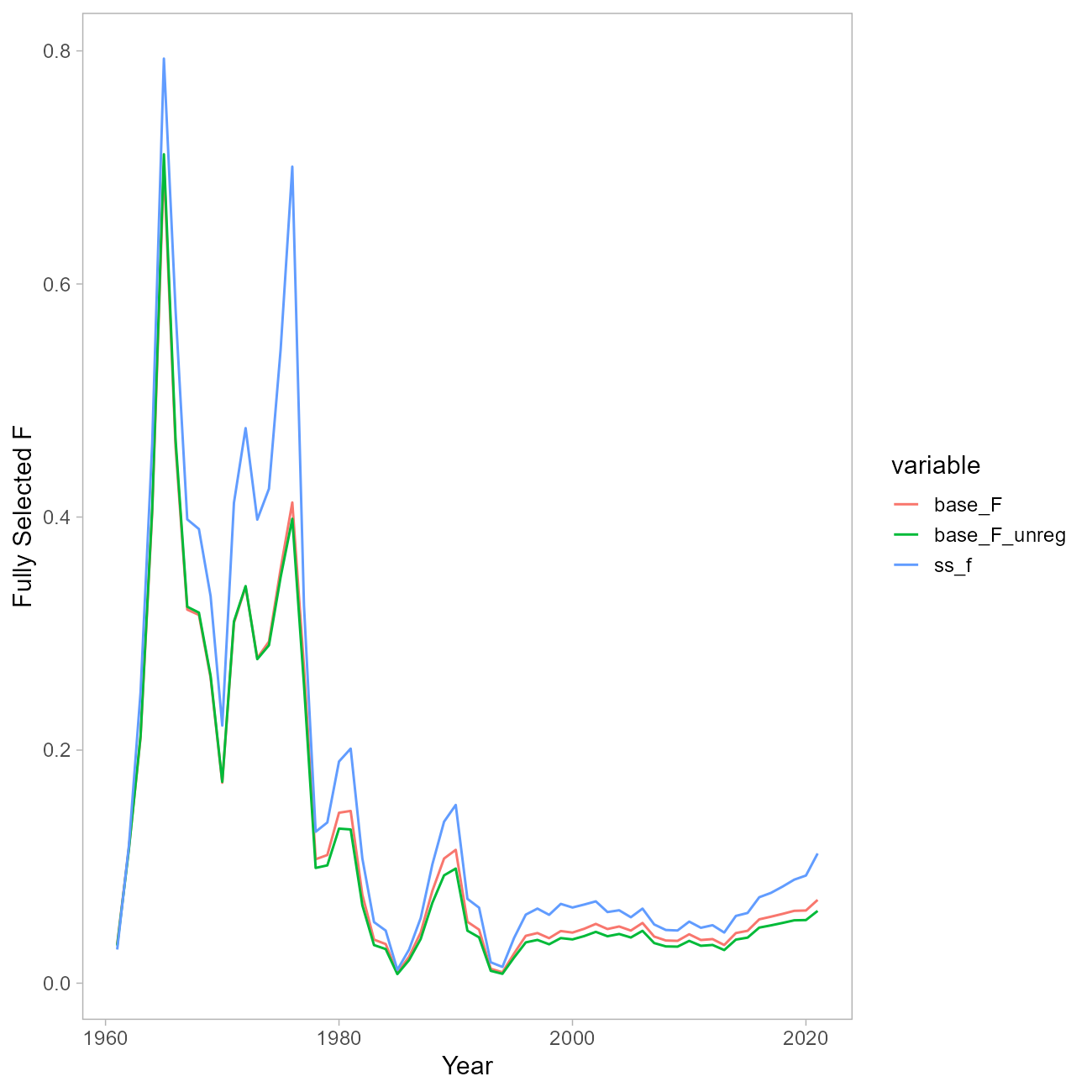


Figure 10.35. Comparison of F trajectories between the base model and a model with the regularization penalty on F disabled, and the SS model.

## References

Methot, R.D., Wetzel, C.R., 2013. Stock synthesis: A biological and statistical framework for fish stock assessment and fishery management. Fisheries Research 142, 86–99. <https://doi.org/10.1016/j.fishres.2012.10.012>

Thorson, J.T., Maureaud, A.A., Frelat, R., Mérigot, B., Bigman, J.S., Friedman, S.T., Palomares, M.L.D., Pinsky, M.L., Price, S.A., Wainwright, P., 2023. Identifying direct and indirect associations among traits by merging phylogenetic comparative methods and structural equation models. Methods Ecol. Evol. n/a. <https://doi.org/10.1111/2041-210X.14076>