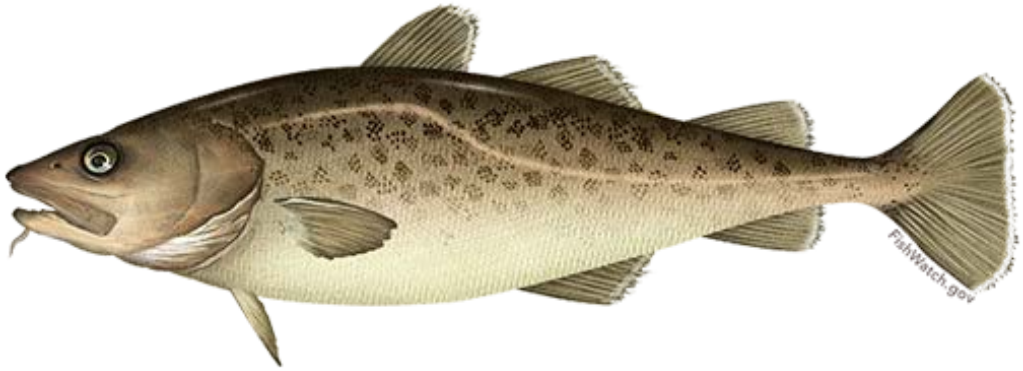
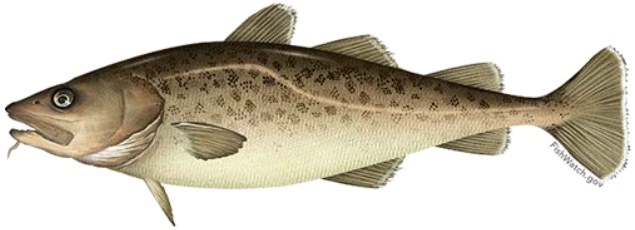
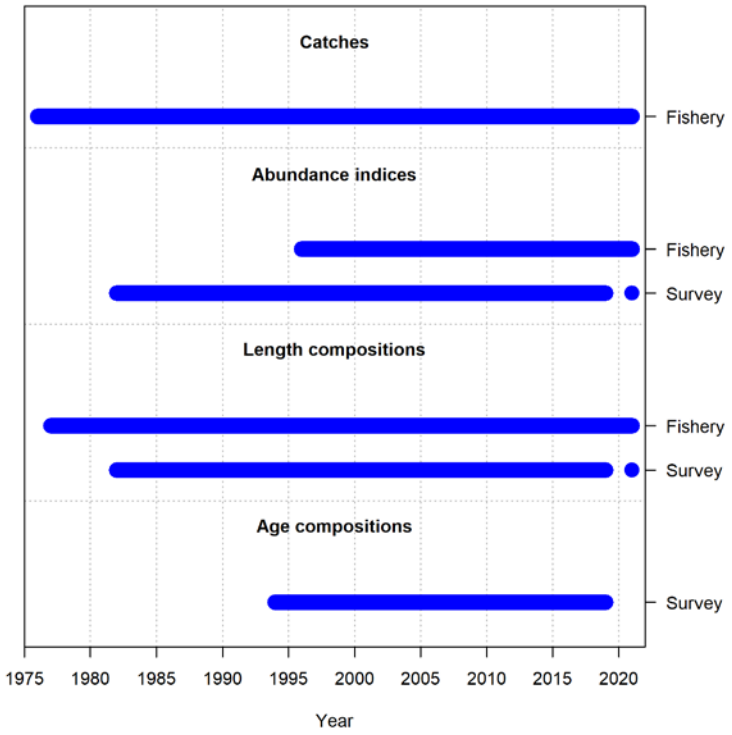
Bering Sea Pacific Cod September Edition

Author: Steve Barbeaux



Bering Sea Pacific cod





- Data
  - VAST survey and CPUE index available
  - VAST Survey age composition
  - Fishery catch weight
  - Fishery and survey size composition
- Models (4 in the ensemble)
  - Stock synthesis age-structured models
  - Single fishery, single season
  - Annually varying survey and fishery size based selectivity
  - Dirichlet multinomial for composition data

Feature	*M19.12a*	*M19.12*	*M20.1*	*M20.2*
Feature 1: Allow catchability to vary?	No	Yes	No	No
Feature 2: Allow domed survey selectivity?	No	No	Yes	No
Feature 3: Use fishery CPUE?	No	No	No	Yes

Explorations for 2022

- New script for the seasonally corrected annual weight at length relationship fit outside the model.
- Removing the seasonally corrected annual weight at length relationship from the model (NOWL).
- New algorithm used for constructing the fishery length composition data using a developed R script.
- Alternative aging bias assuming bias in those otoliths aged prior to 2007 and no bias in those aged after 2007 instead of bias assumed in 1994-2007 and 2008+ blocks. (AGE)
- Alternative input sample size used for the fishery length composition and additional tuning to ensure the Dirichlet multinomial log theta parameter is not fit at or near a bound. (WT)
- Fitting an additional standard error term on the VAST bottom trawl survey index. (SE)

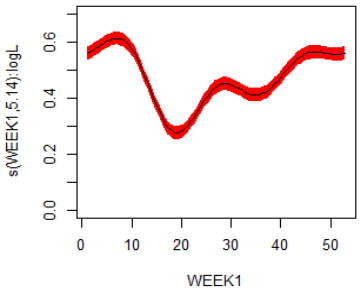
Seasonally corrected annual weight at length relationship

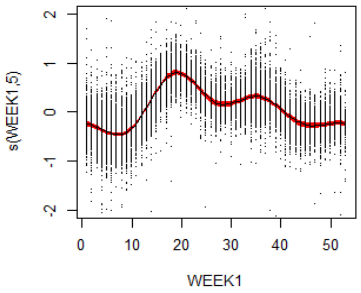
- Initially developed by Thompson to deal with seasonal and annual changes in weight at length in the fishery data.
  - Linear model in a now unsupported version of MathCad.
    - Replicating this effort in MathCad was no longer feasible

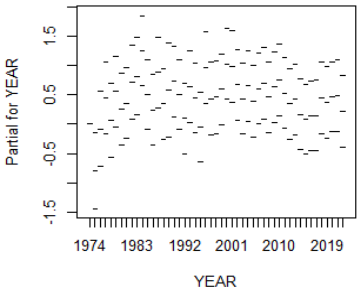
- Barbeaux replicated effort in R using a generalized additive model
  - Similar results, much simpler code.

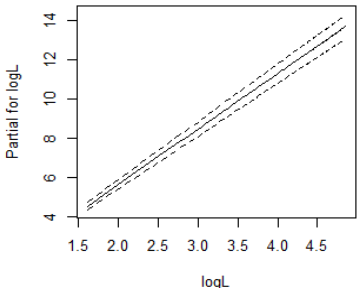
Seasonally corrected annual weight at length relationship

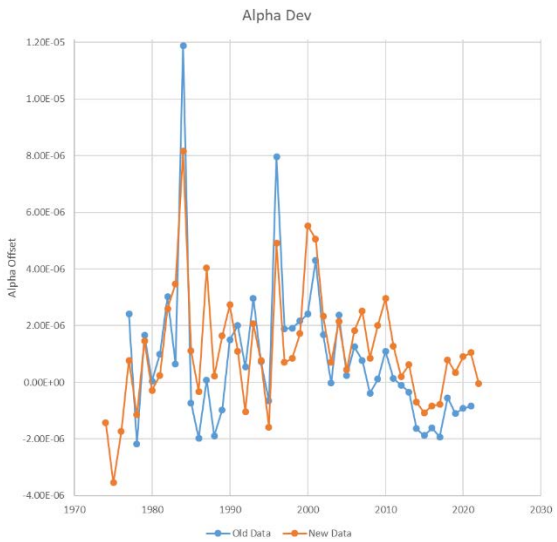
- GAM for predicting weekly variability on growth
- The annual mean weight at length is then calculated from the linear growth model
- Results are an index of annual residuals on alpha and beta for the weight at length relationship

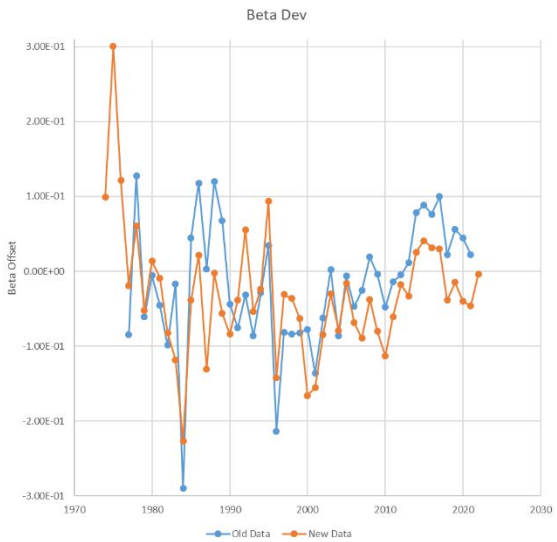






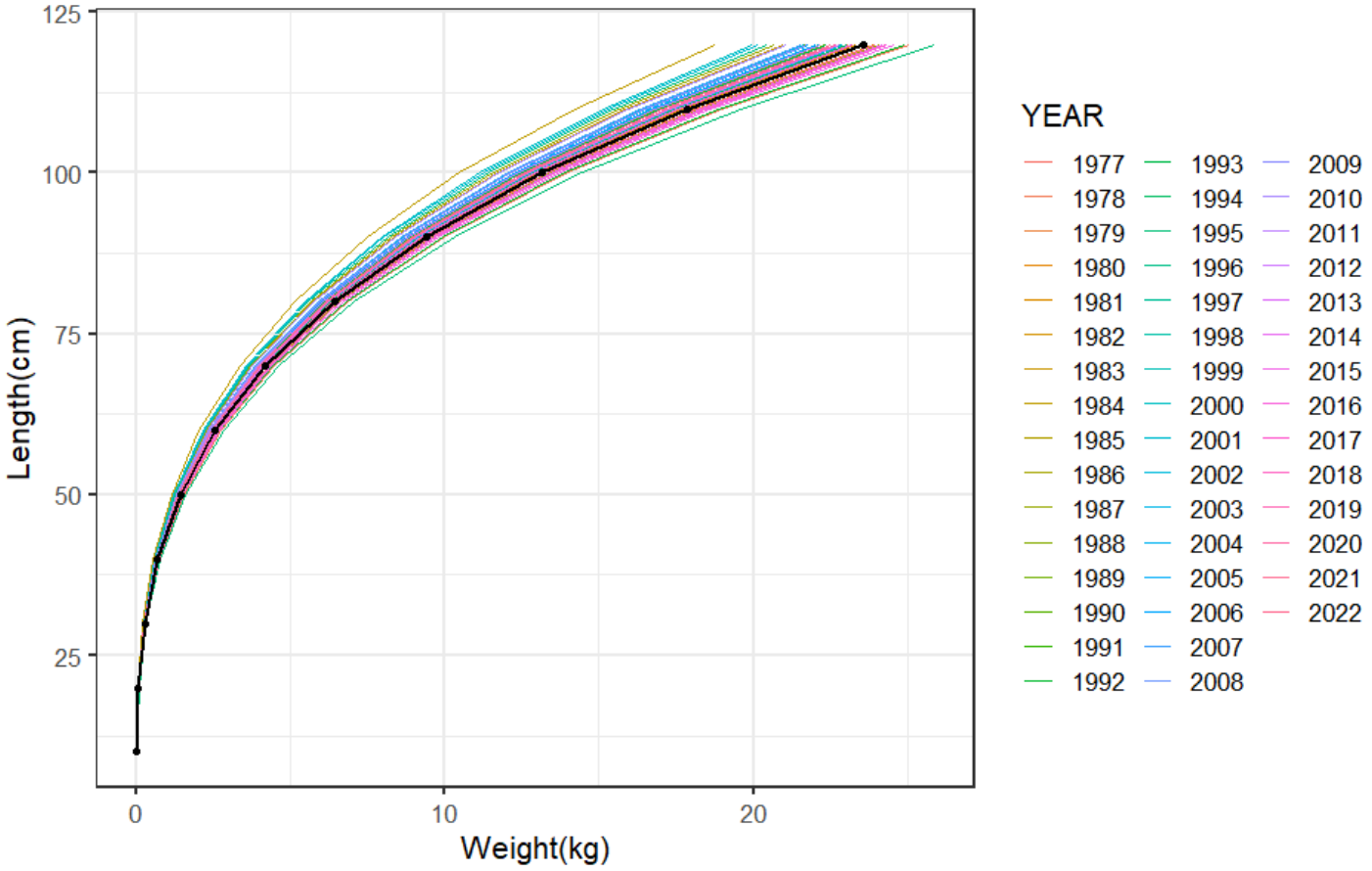




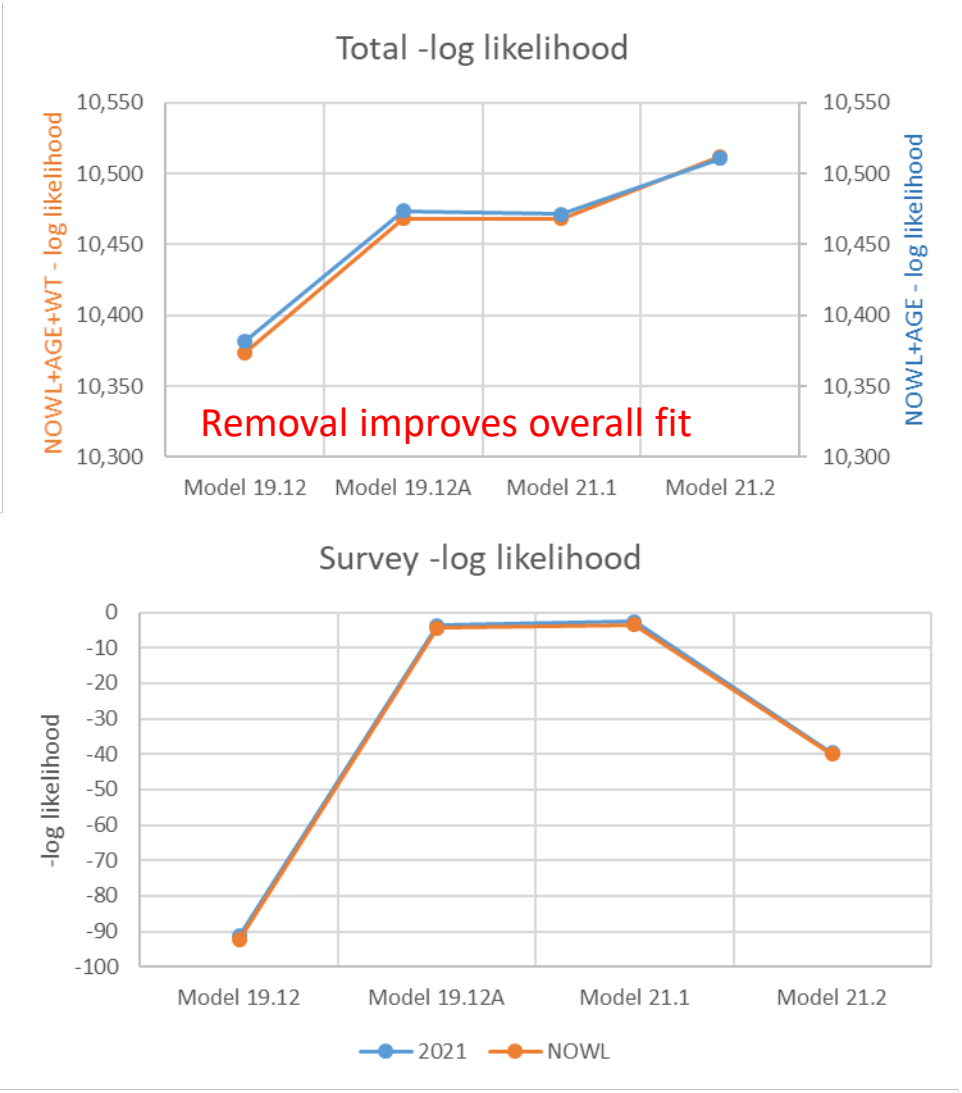


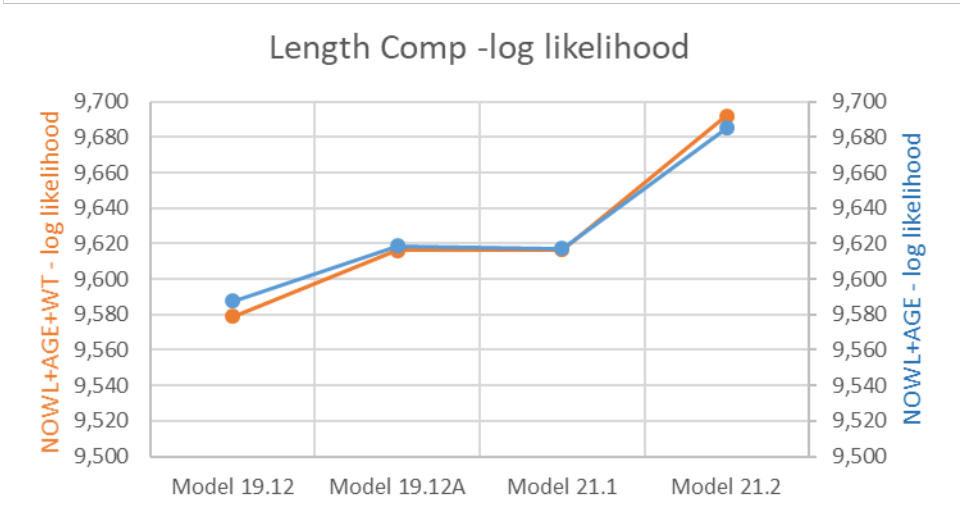
Note that the results are similar to Thompson model, but not exact

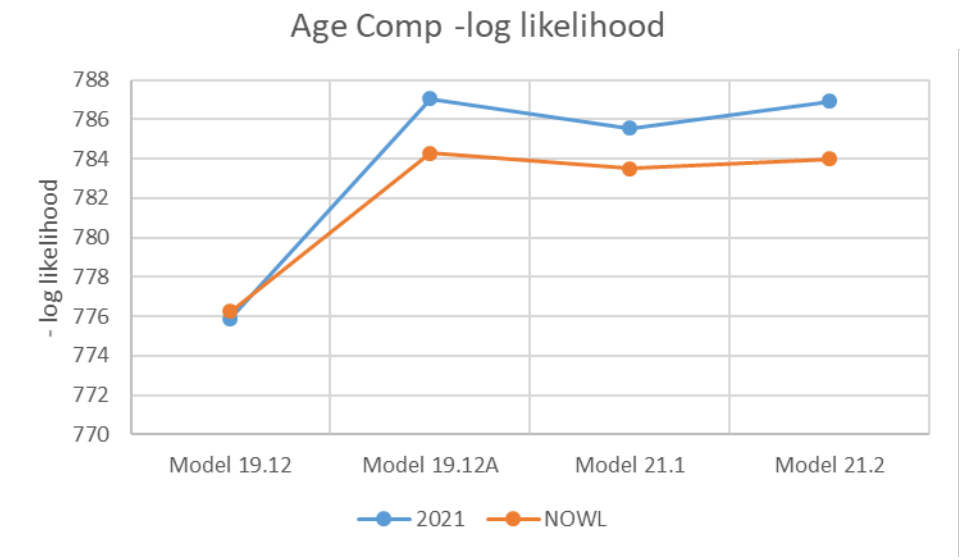
Seasonally corrected annual weight at length relationship



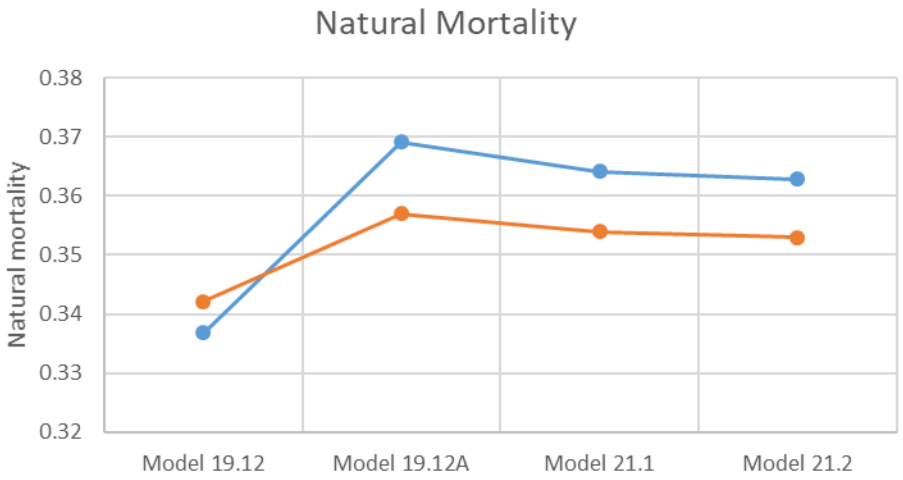
Removing the seasonally corrected weight at length relationship (NOWL)

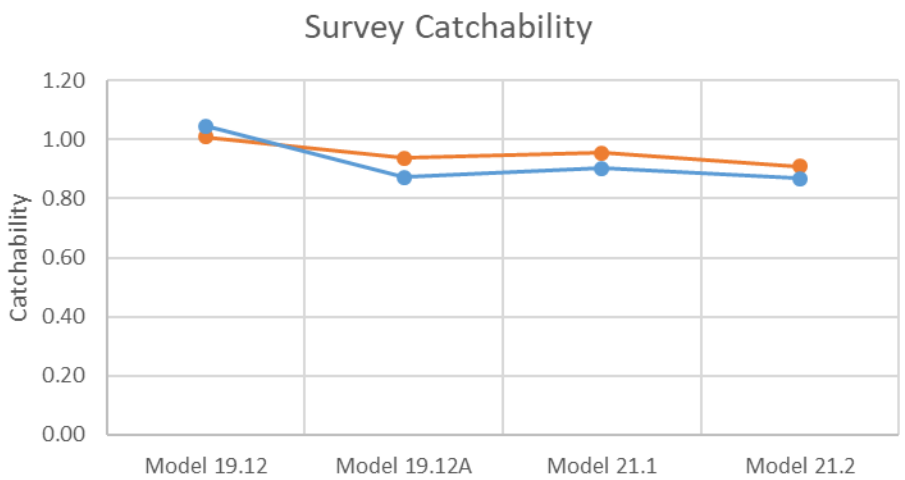


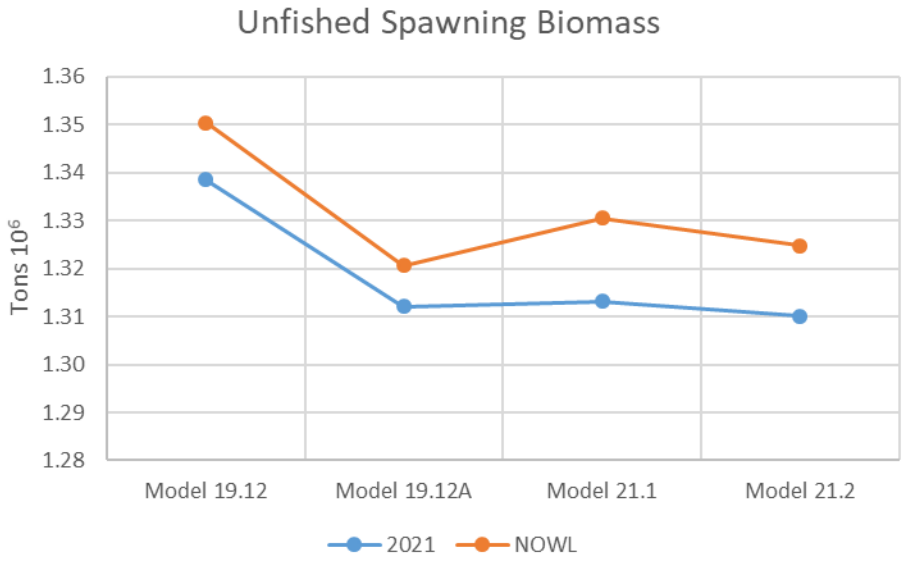


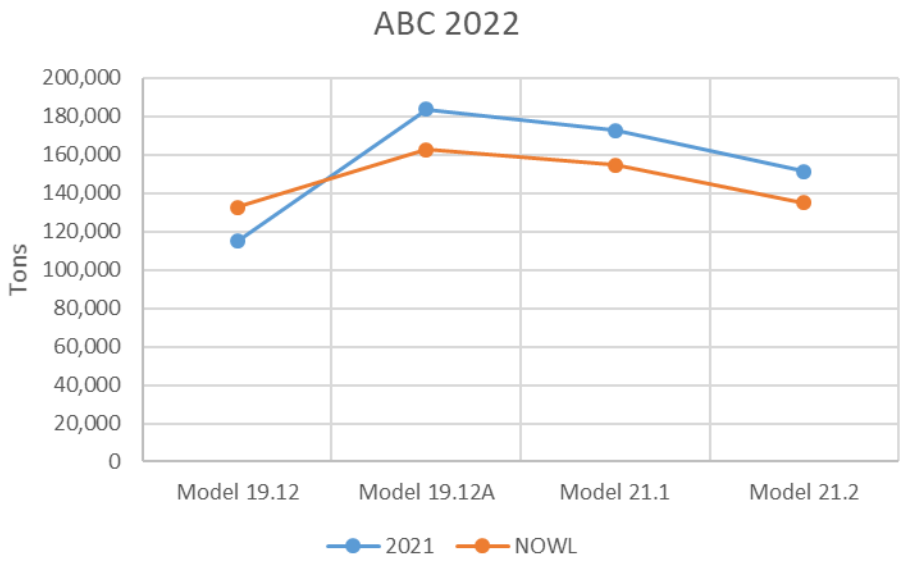


Removing the seasonally corrected weight at length relationship (NOWL)

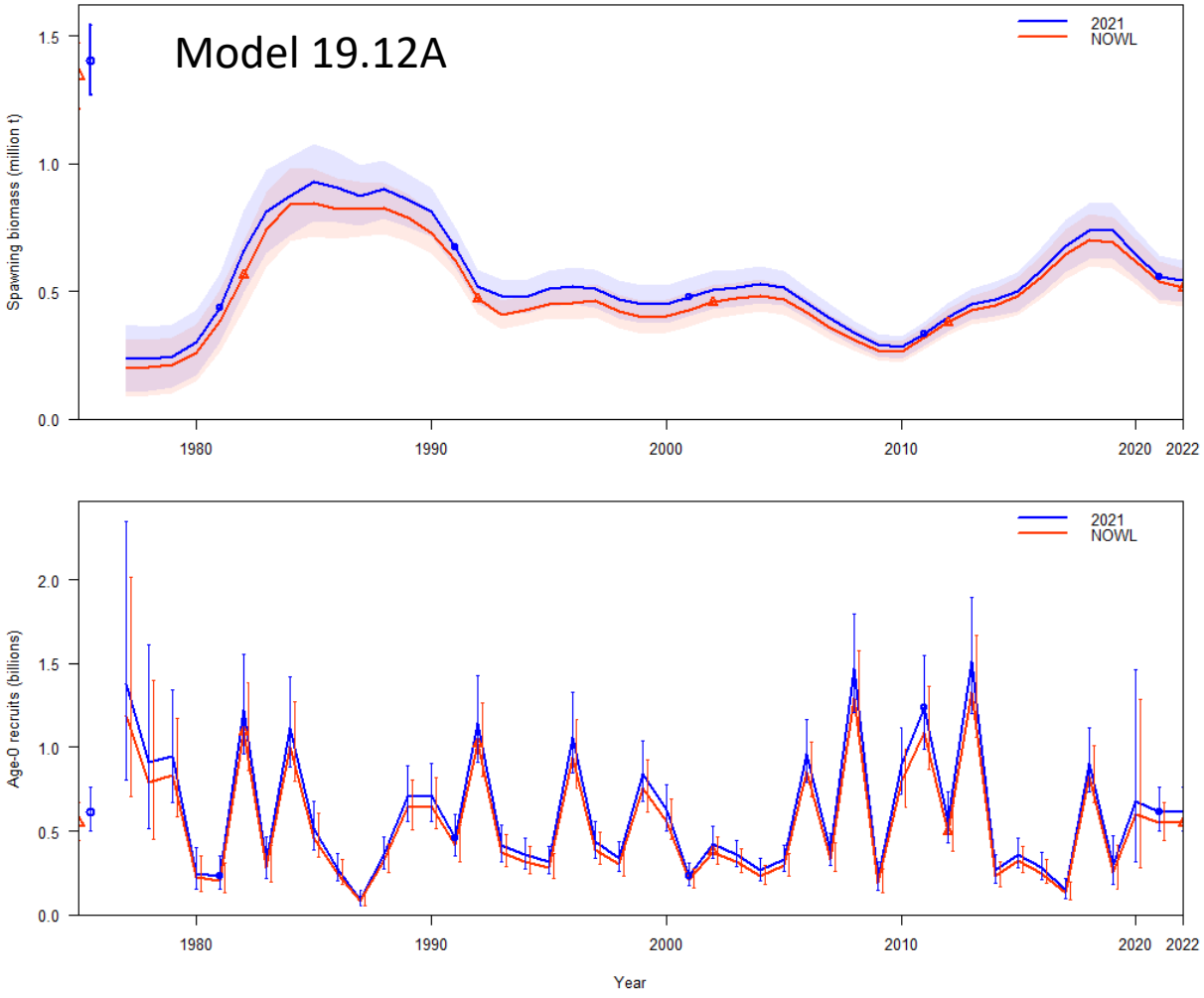








Removing the seasonally corrected weight at length relationship (NOWL)



Minor changes in natural mortality and catchability lead to minor decrease in spawning biomass and recruitment in all models

Author Recommendation

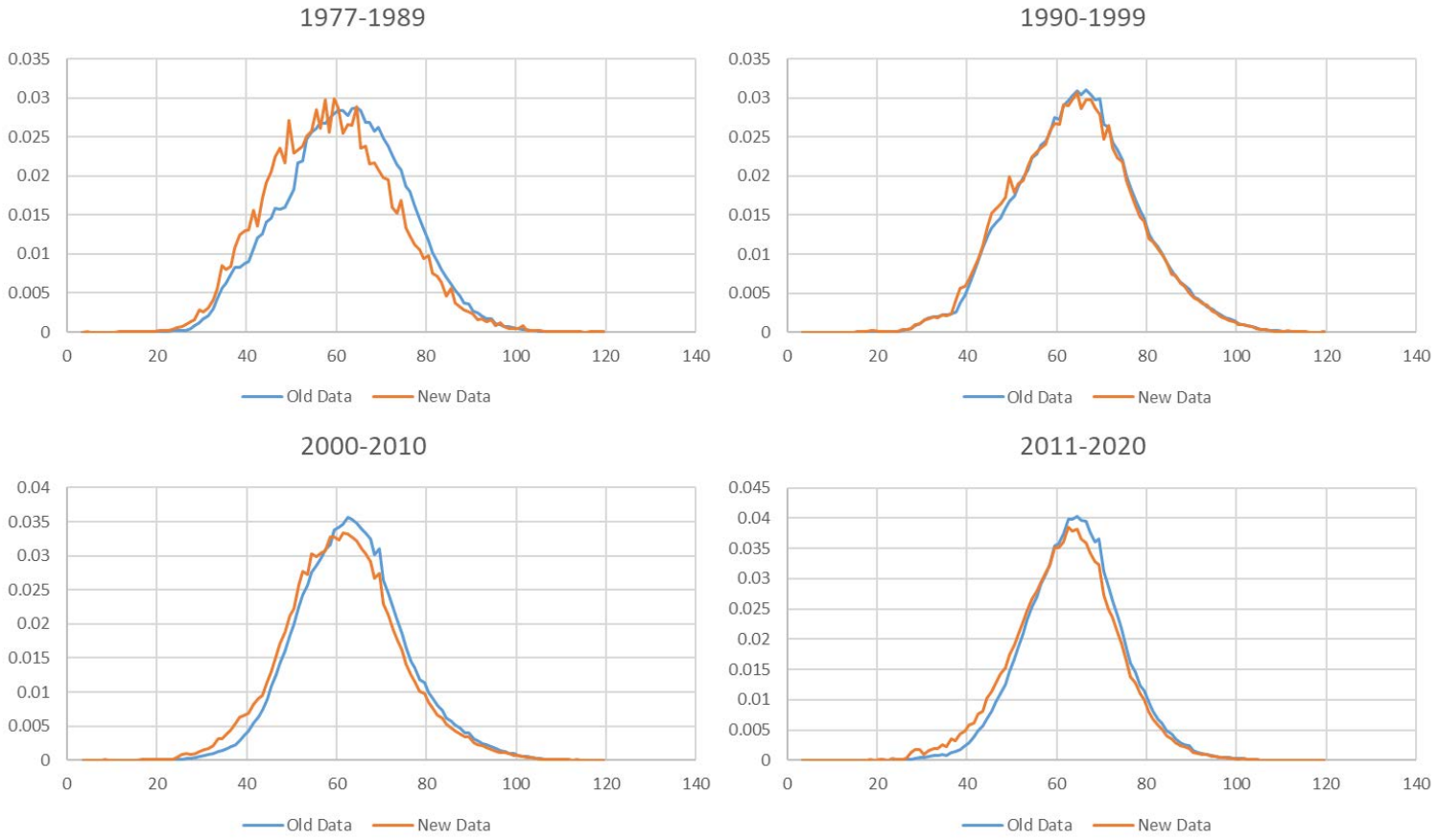
Because of the lack of improvement to fit by including it and difficulty in projecting this relationship, I recommend that the seasonally corrected annual weight at length relationship used in the base model be discarded for 2022 and that we explore other options for modeling seasonality and annual changes in growth in 2023.

Change in fishery length composition data

- Thompson used catch weight by month, area, gear, and year to weight length compositions 1989-2021
  - All processing conducted in excel
  - 1977-1988 unweighted length compositions were used

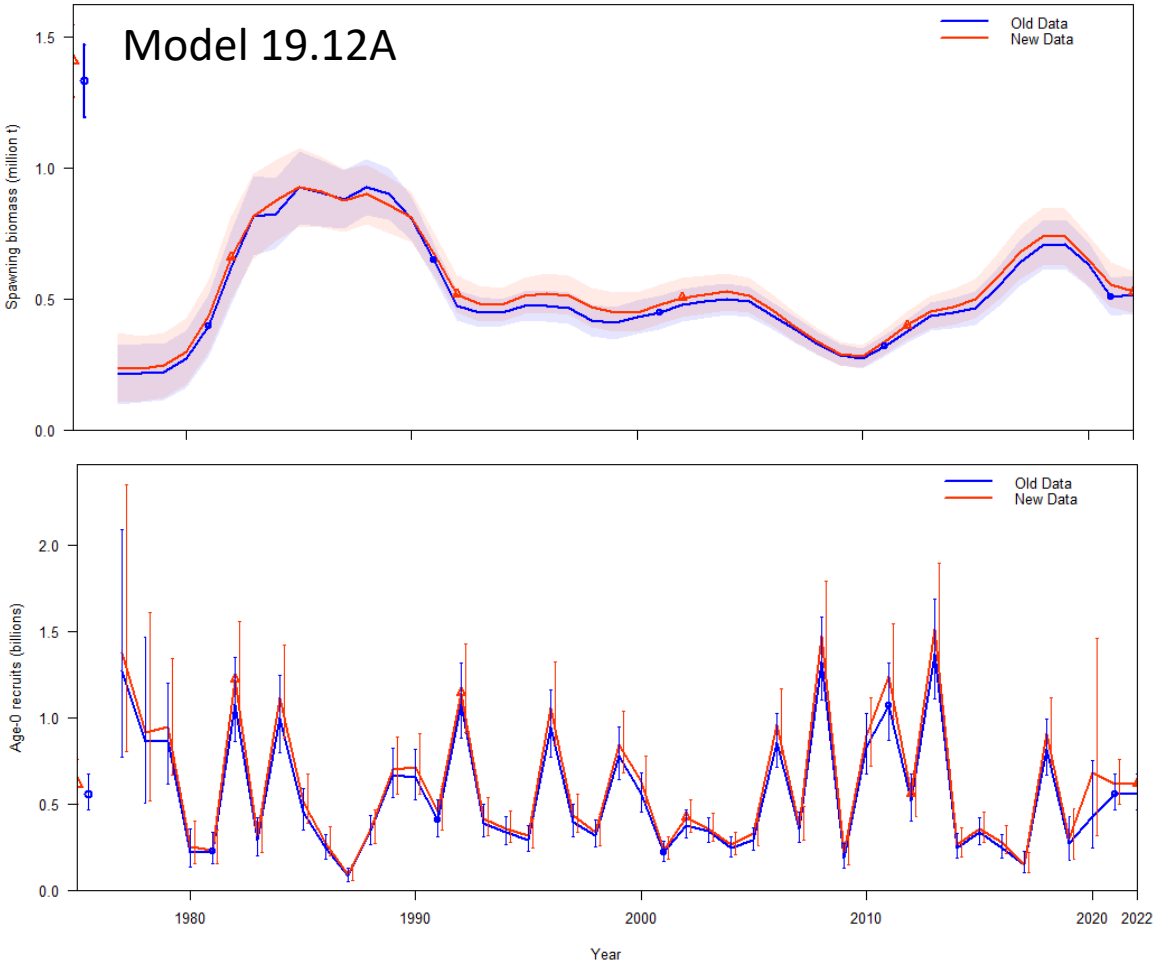
- New method used catch number by haul/set, month, area, gear, and year to weight length composition samples
  - All processing conducted in R and documented in an R function.
  - All years weighted the same

Change in fishery length composition data



Small shift to smaller fish in the new fishery length composition method

Change in model results due to data changes



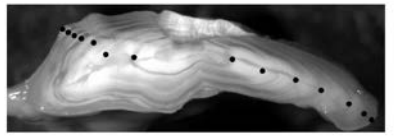
Minor changes in natural mortality and catchability leading to minor increase in spawning biomass and recruitment

Author Recommendation

For ease of use and more appropriate weighting of older data I recommend the new R-script process developed for producing the fishery length composition data be used.

Model Changes Explored

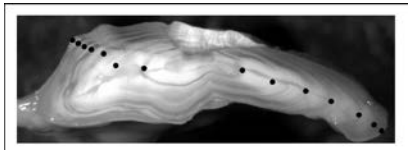
- **+AGE** = Aging bias changed from two blocks (1977-2007 and 2008+) to a single block (1977-2007)
- **+WT** = Changing input sample sizes for fishery and survey length composition data
- **+SE** = Fitting additional standard errors to abundance indices



Alternate aging bias (+AGE)

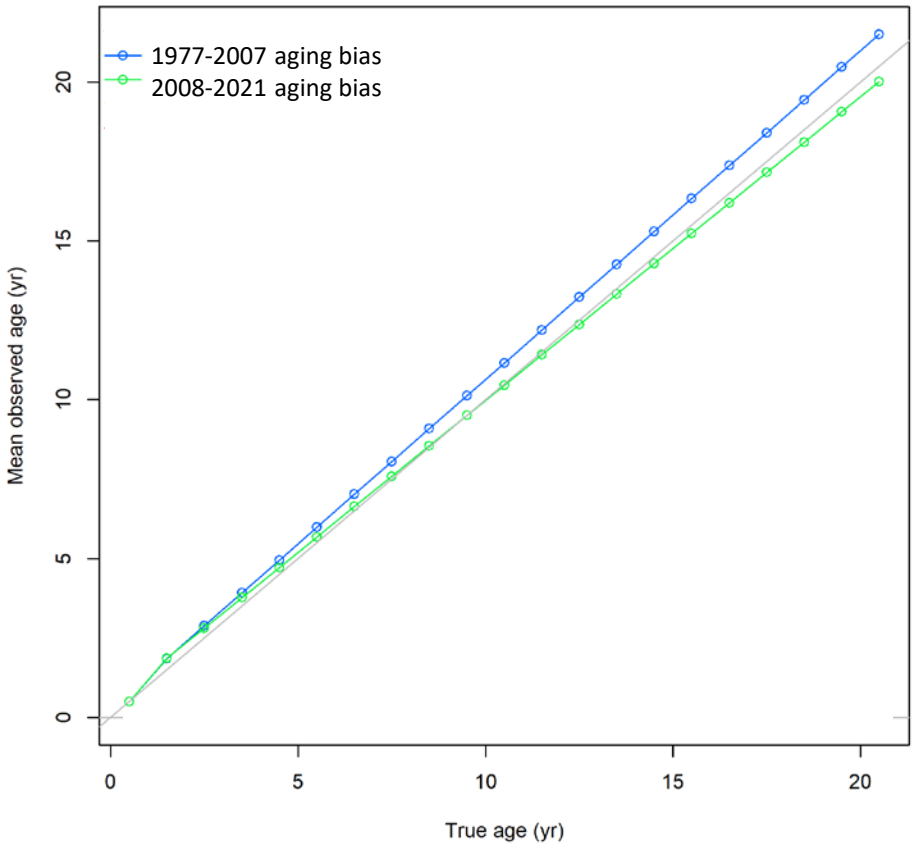
- Thompson models have two blocks for aging bias
  - 1992-2007 positive bias with over aging
  - 2008-2021 negative bias with under aging

- Explored models assume one block for aging
  - 1992-2007 positive bias with over aging
  - 1992-2007 aging bias confirmed through isotope analysis Kastelle et al. (2016)
  - New aging methods assumed unbiased

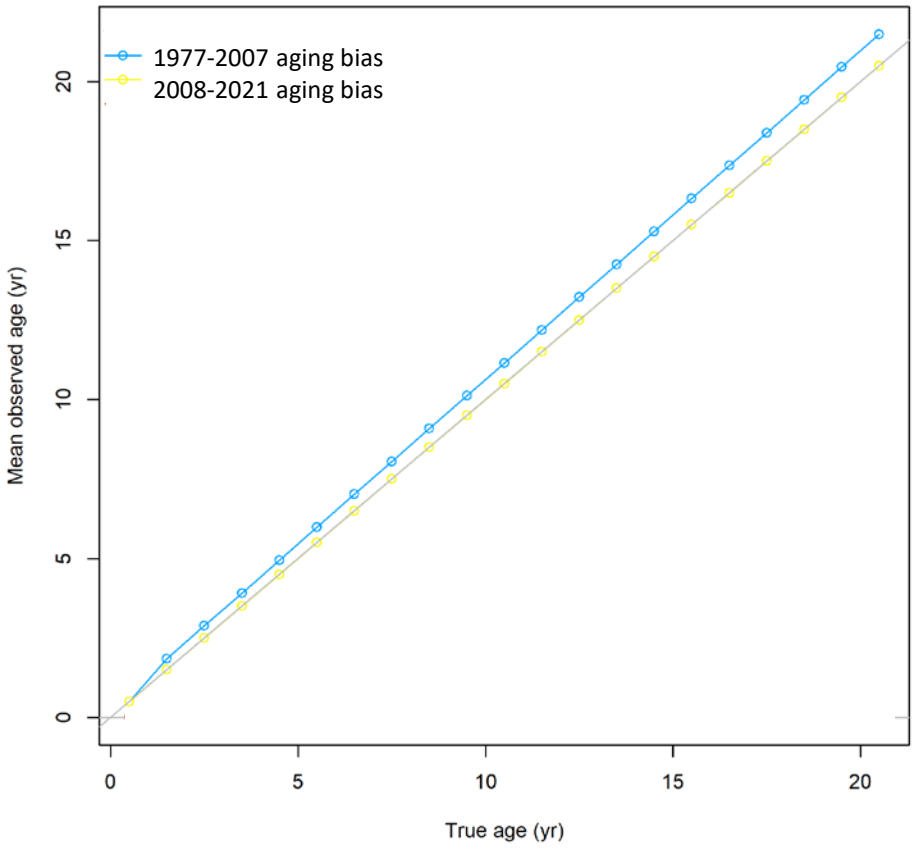


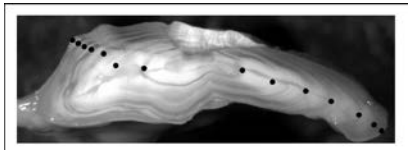
Alternate aging bias (+AGE)

2021 models

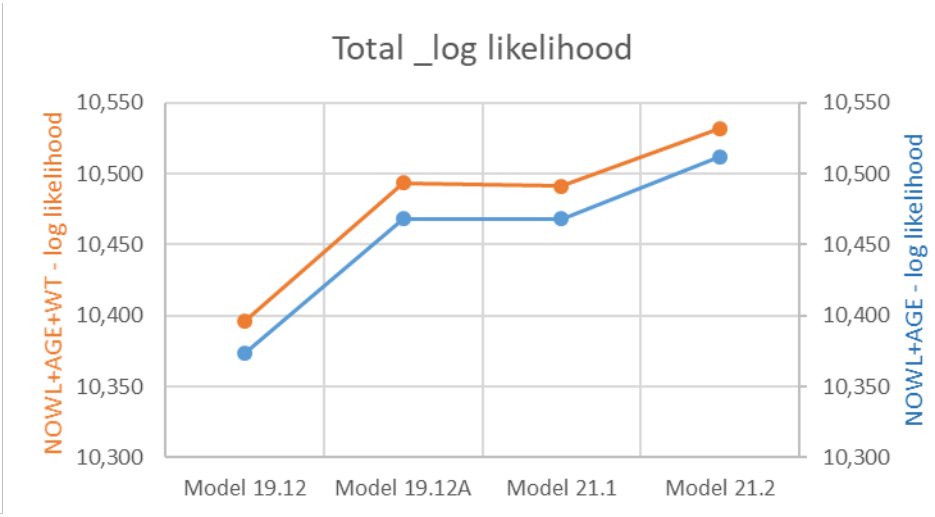


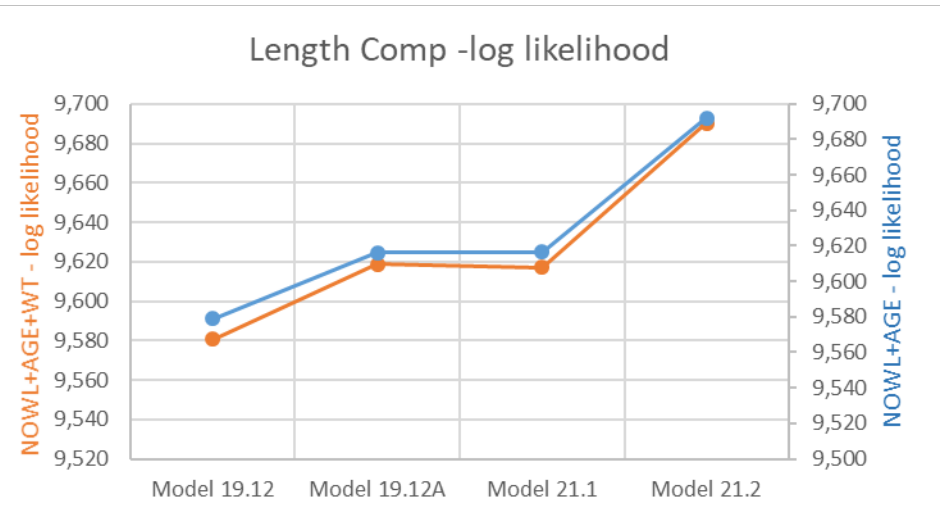
+AGE models

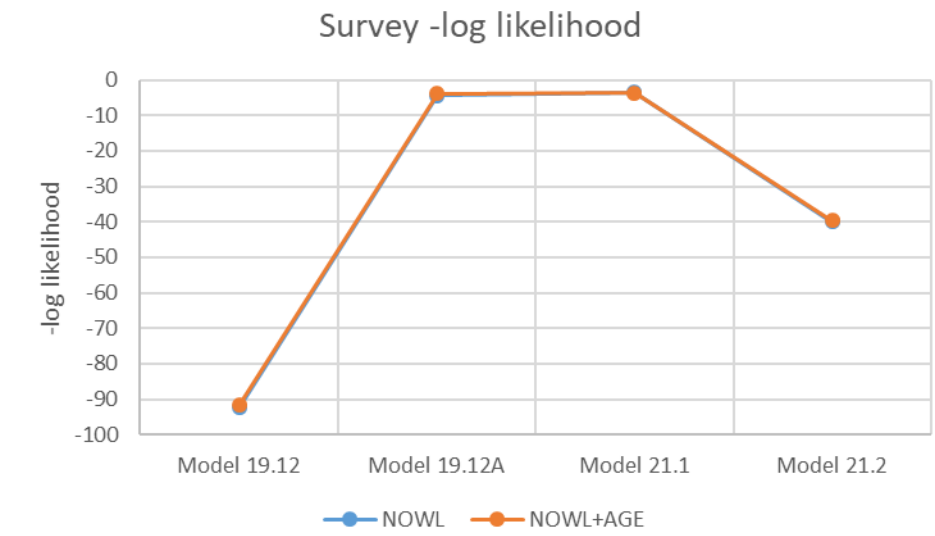


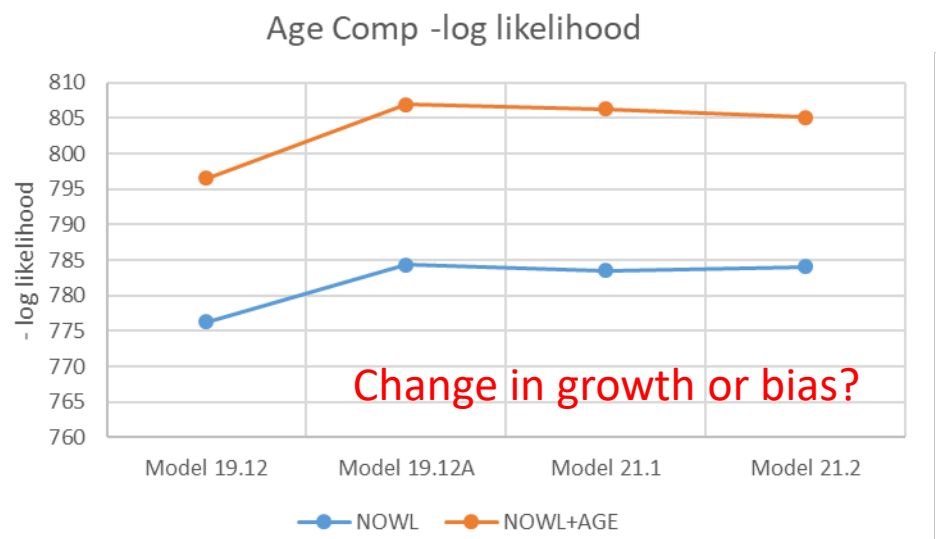


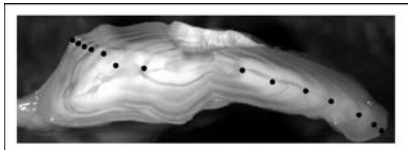
Alternate aging bias (+AGE)



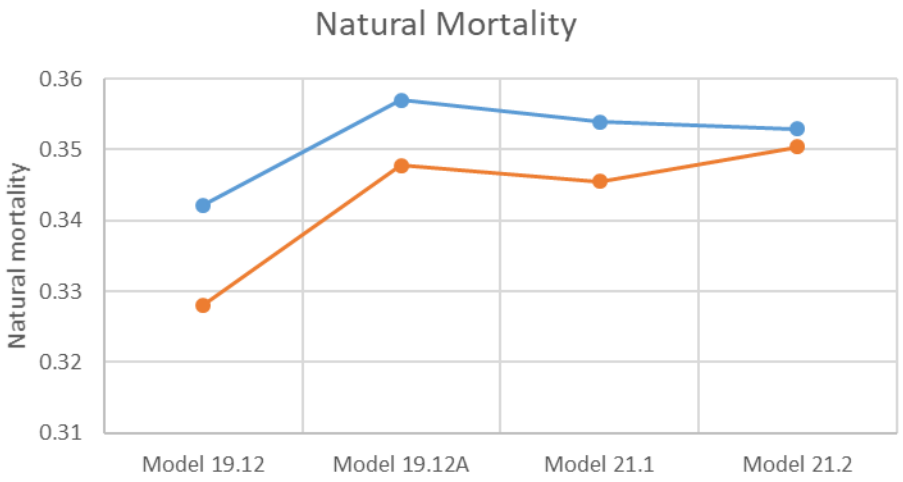


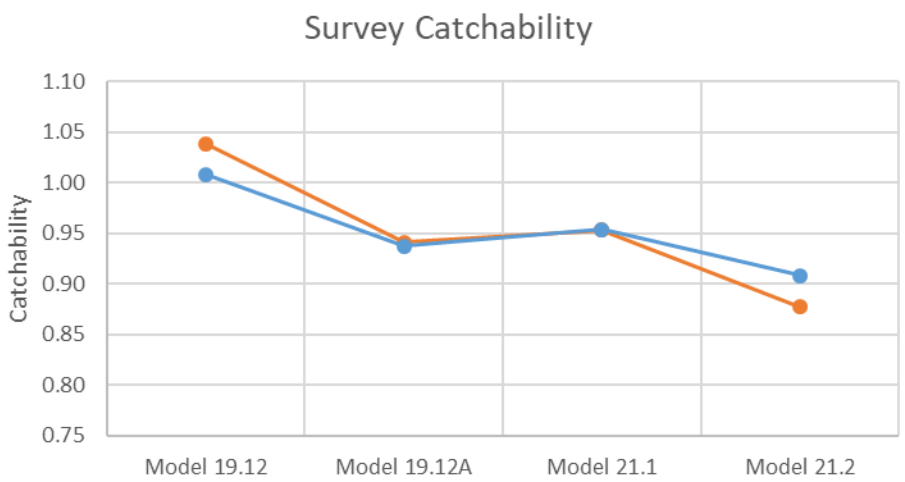


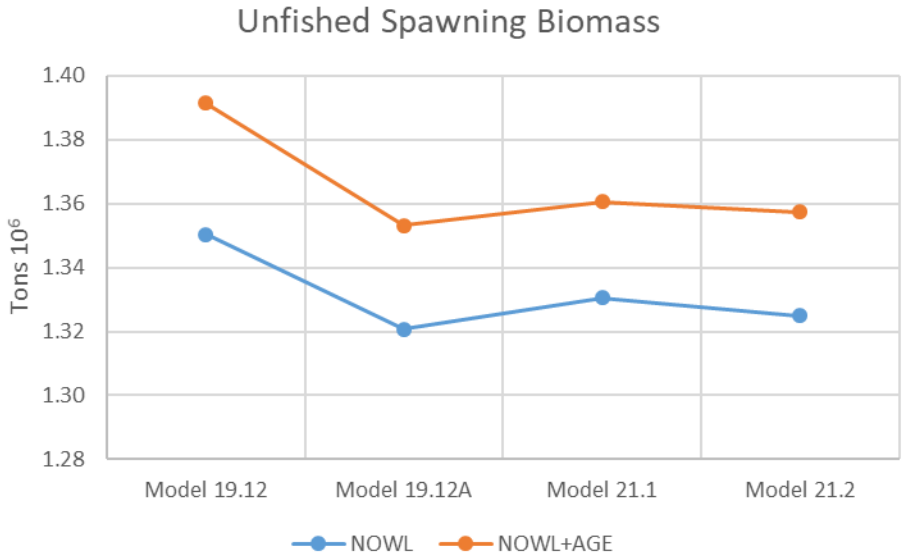


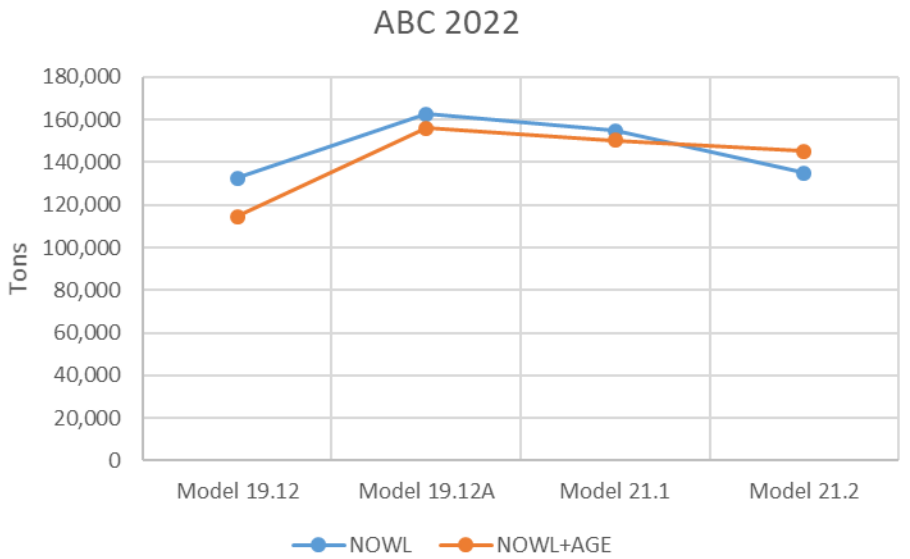


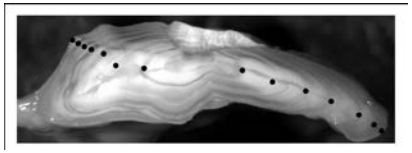
Alternate aging bias (+AGE)



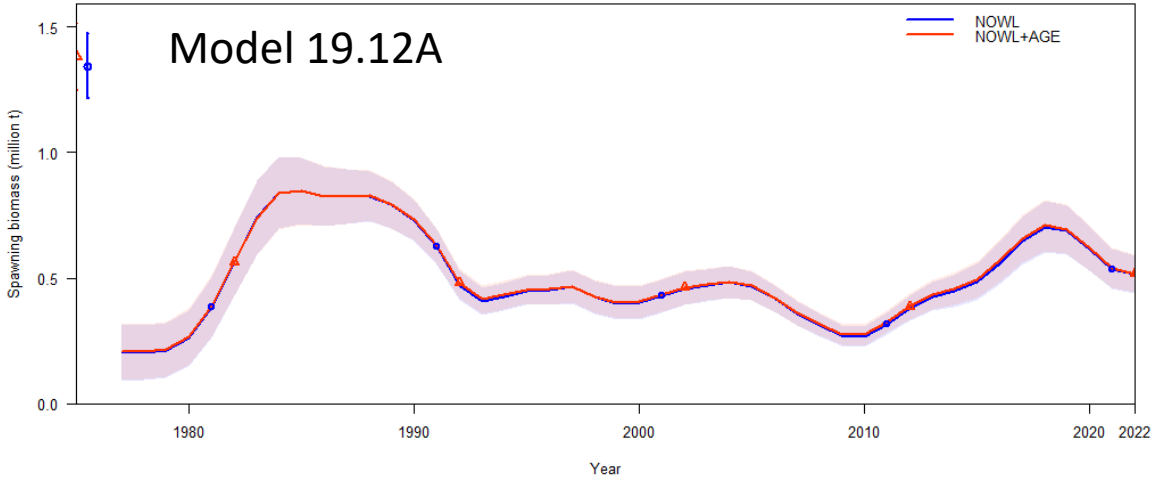


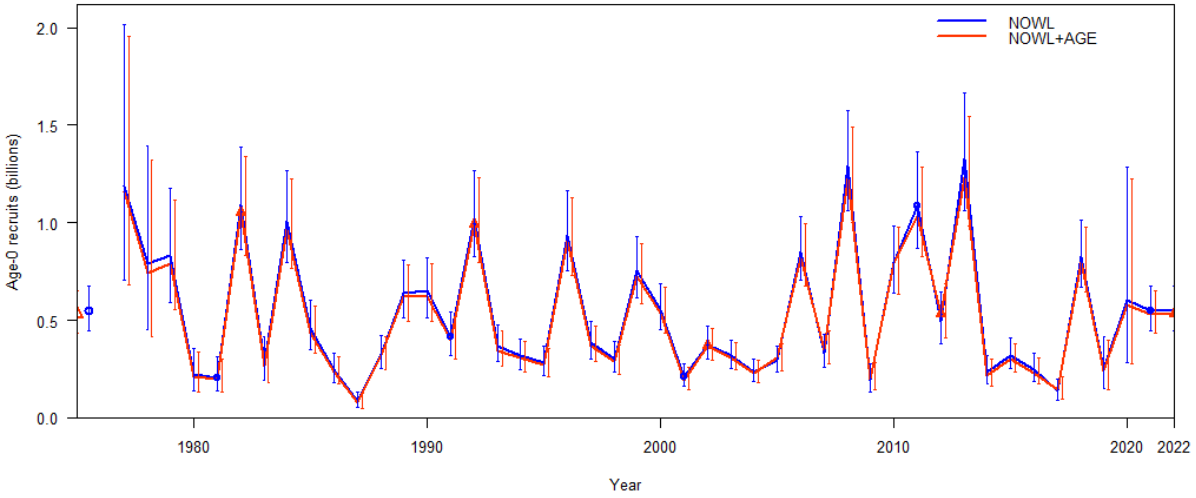






Alternate aging bias (+AGE)





Minor changes in natural mortality leading to very minor decrease in spawning biomass and recruitment

Author Recommendation

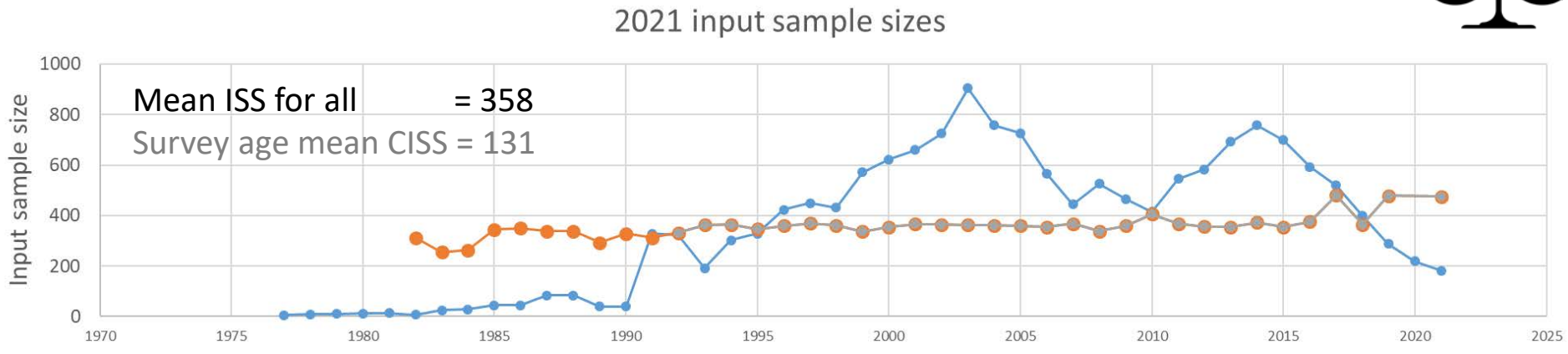
In regards to advice from the Age and Growth Laboratory and despite the degradation in model fit, I recommend that fitting aging bias for the most recent time period be removed for the 2022 models and that I explore more options for capturing variability in growth in 2023.

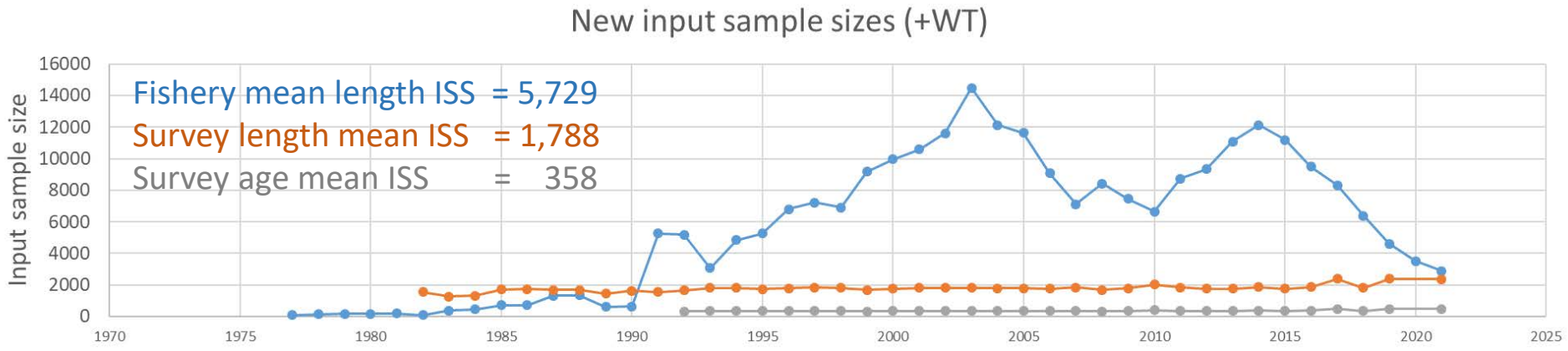
New input sample sizes (+WT)

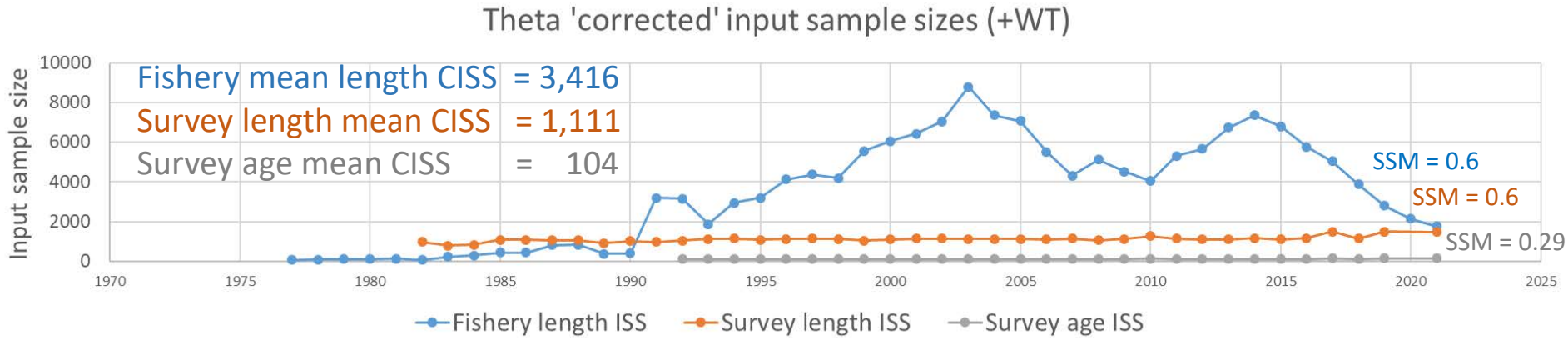
- Thompson
  - Reduced mean fishery length composition input sample size to the mean number of hauls in the bottom trawl surveys
    - Mean input sample size for all = 358
    - Dirichlet multinomial log theta parameter is at the upper bound for both survey and fishery length composition data.

- +WT
  - Fishery length composition input sample size
    - Number of hauls sampled for lengths
    - Mean input sample size = 5,729
  - Survey length composition input sample size
    - Number of survey hauls Increased iteratively until log theta fit off bounds
    - Mean input sample size for survey length comps (X5) = 1,788
  - Survey Age comps input sample size
    - Number of survey hauls
    - Mean input sample size for survey age comps = 358

Model 19.12A







CISS = Dirichlet 'corrected' input sample size

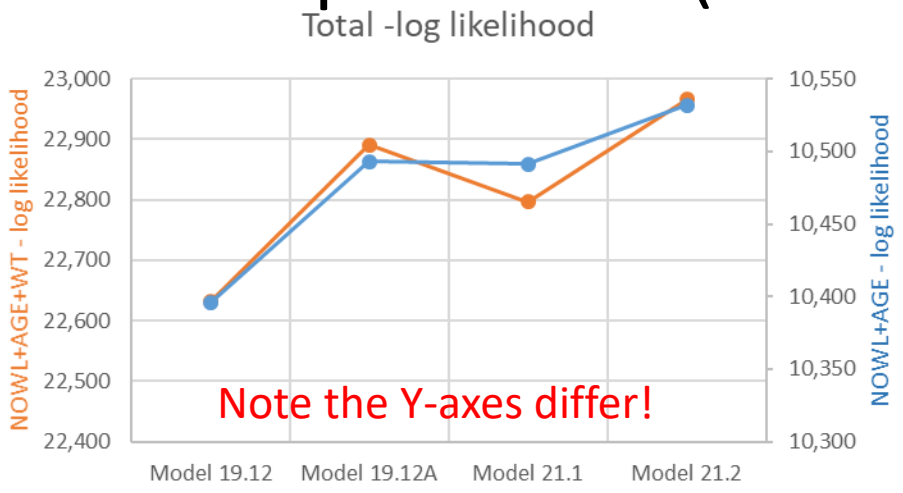
SSM = Dirichlet Sample Size Multiplier

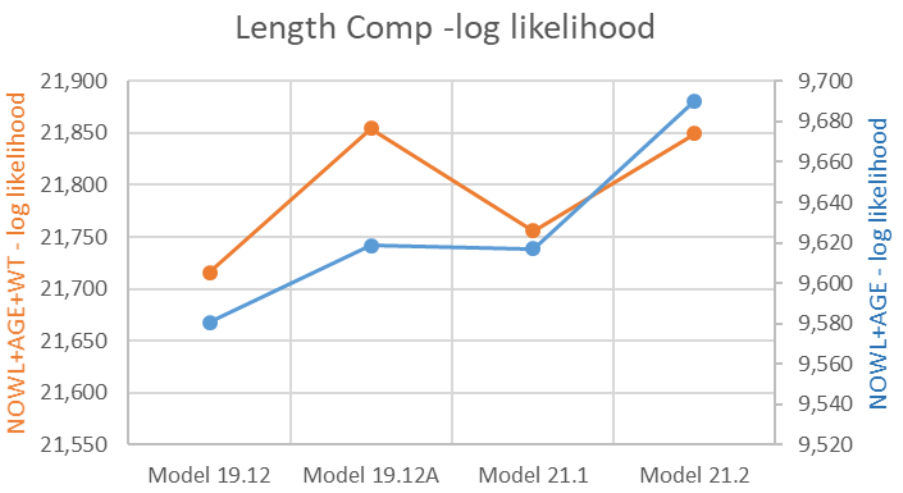
Model changes due to input sample size (+WT)

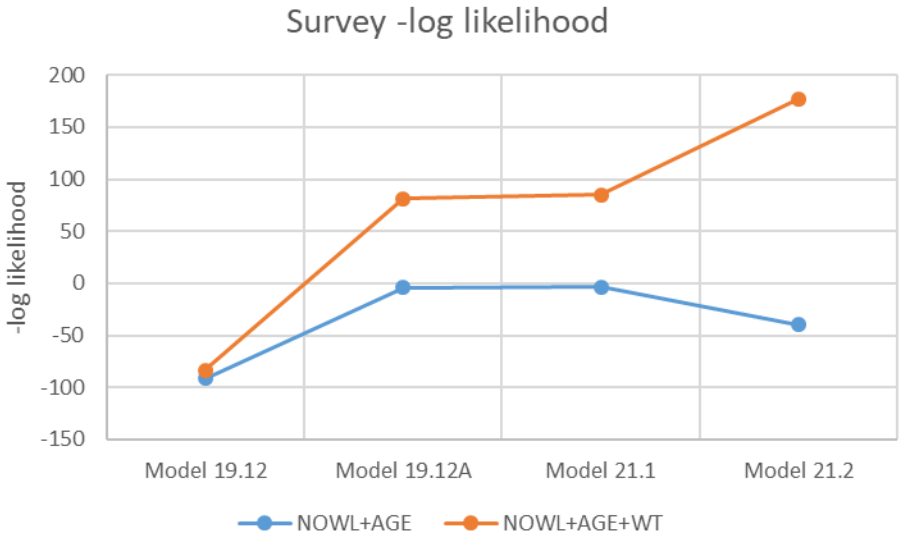
Sample size multiplier	*Model 19.12*	*Model 19.12A*	*Model 21.1*	*Model 21.2*	
Fishery Length	1	1	1	1	NOWL+AGE
Fishery Length	0.643	0.607	0.658	0.633	NOWL+AGE+WT
Survey Length	1	1	1	1	NOWL+AGE
Survey Length	0.589	0.622	0.578	0.547	NOWL+AGE+WT
Survey Age	0.394	0.366	0.371	0.324	NOWL+AGE
Survey Age	0.249	0.290	0.250	0.235	NOWL+AGE+WT

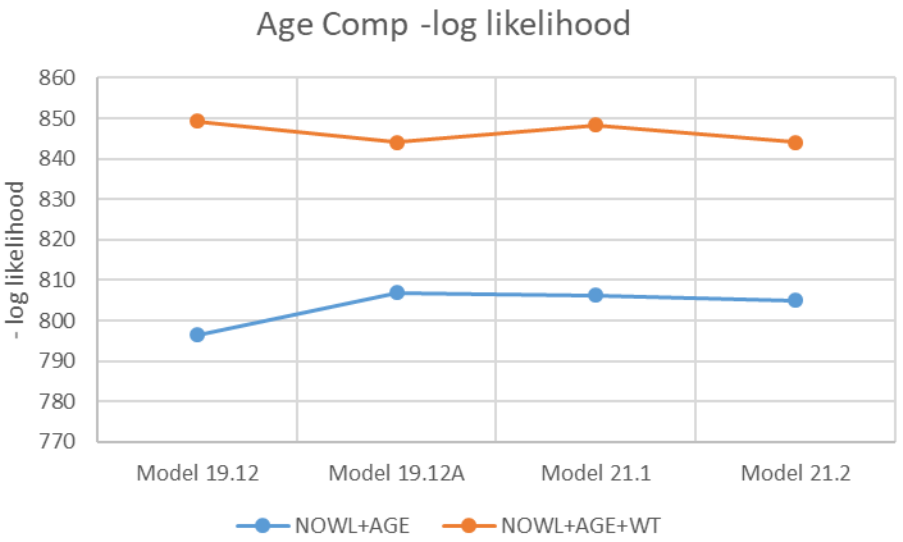
Input sample size	*Model 19.12*	*Model 19.12A*	*Model 21.1*	*Model 21.2*	
Fishery Length	358	358	358	358	NOWL+AGE
Fishery Length	3616	3416	3701	3560	NOWL+AGE+WT
Survey Length	358	358	358	358	NOWL+AGE
Survey Length	1054	1111	1033	979	NOWL+AGE+WT
Survey Age	141	131	133	116	NOWL+AGE
Survey Age	89	104	89	84	NOWL+AGE+WT

Model changes due to input sample size (+WT)



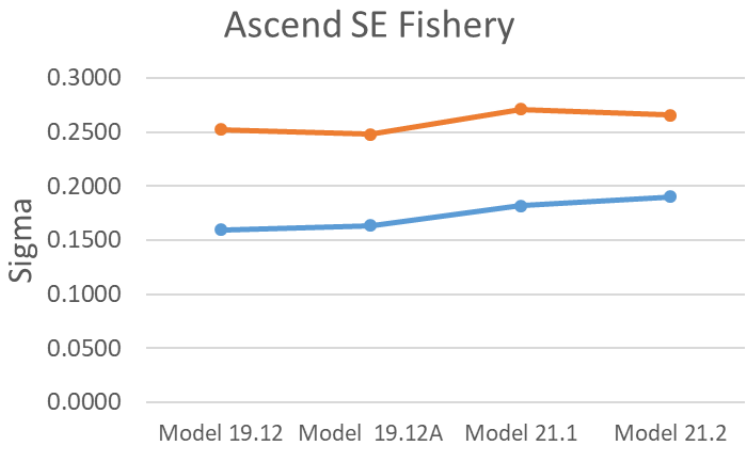


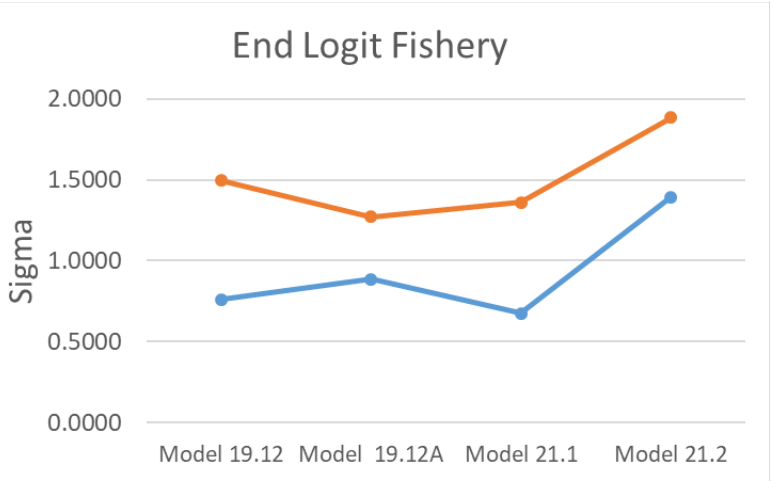


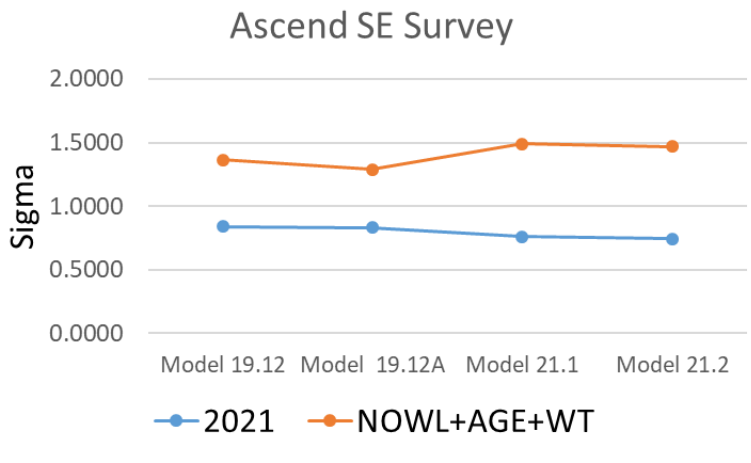


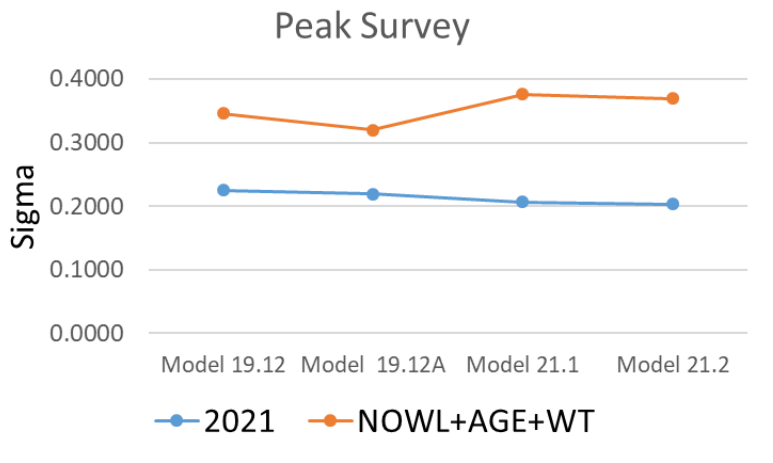
Model changes due to input sample size (+WT)

- Increase in sigma for selectivity parameters







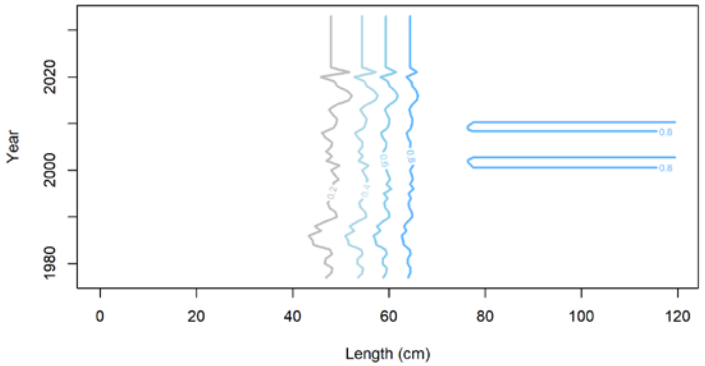


Model changes due to input sample size (+WT)

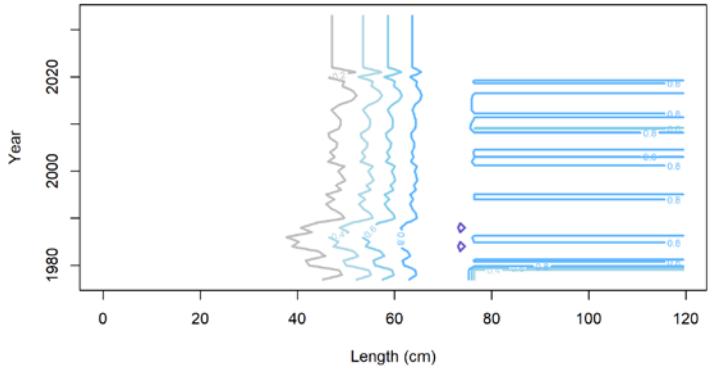
- Increase in sigma for selectivity parameters

Model 19.12A

Fishery

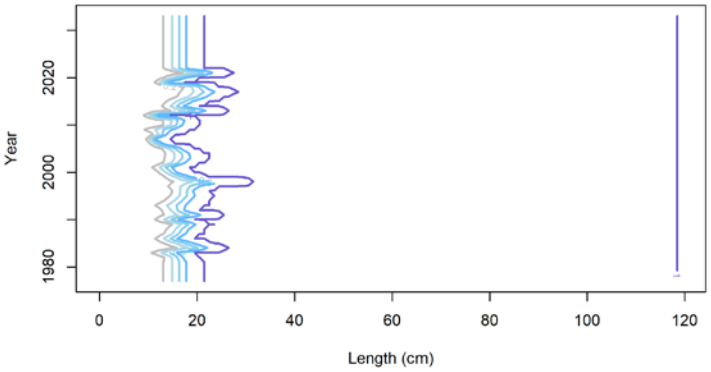


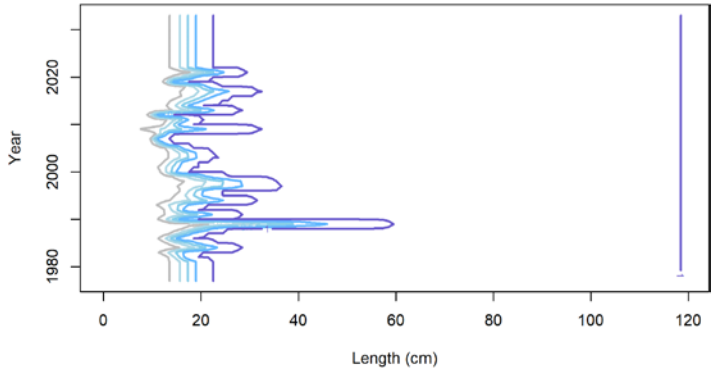
NOWL+AGE



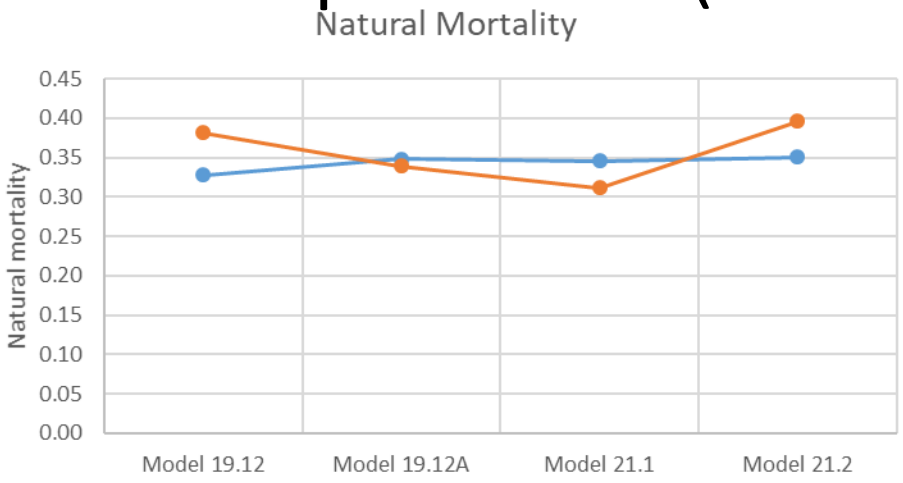
NOWL+AGE+WT

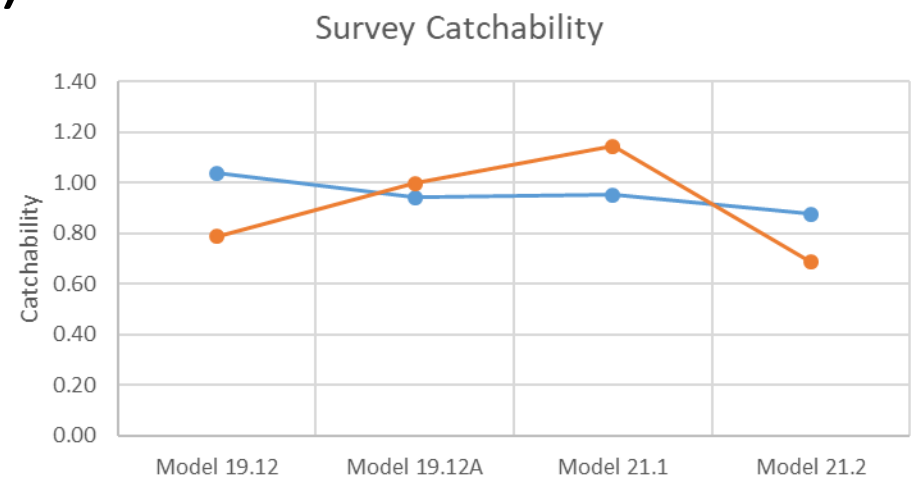
Survey

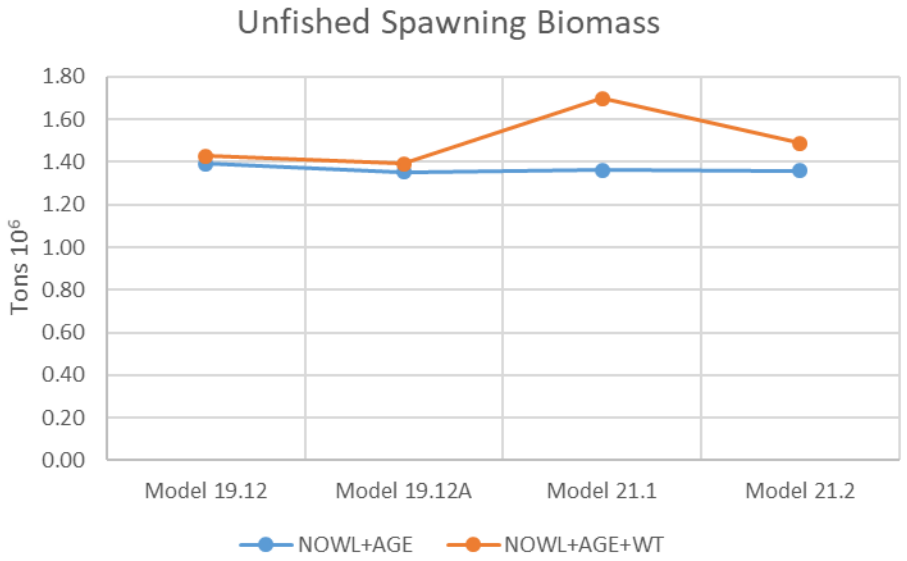


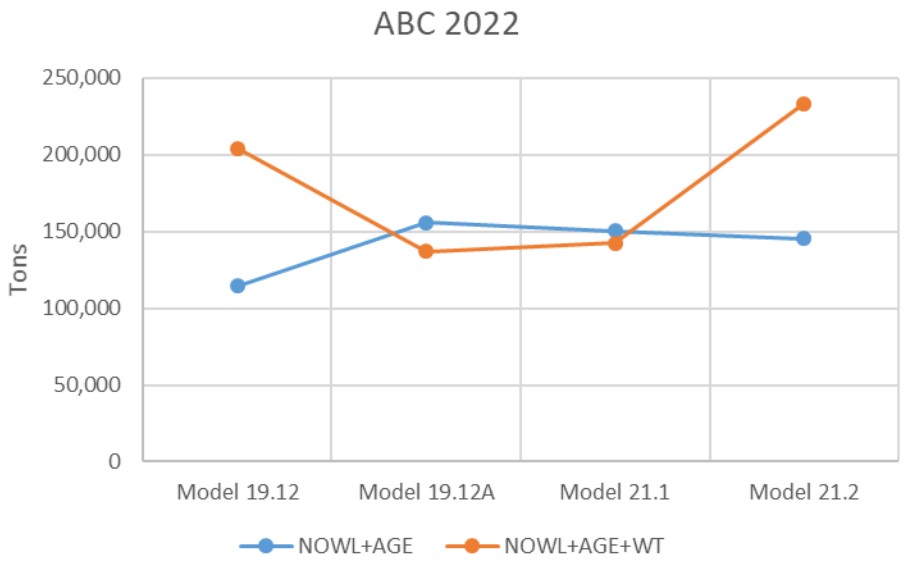


Model changes due to input sample size (+WT)





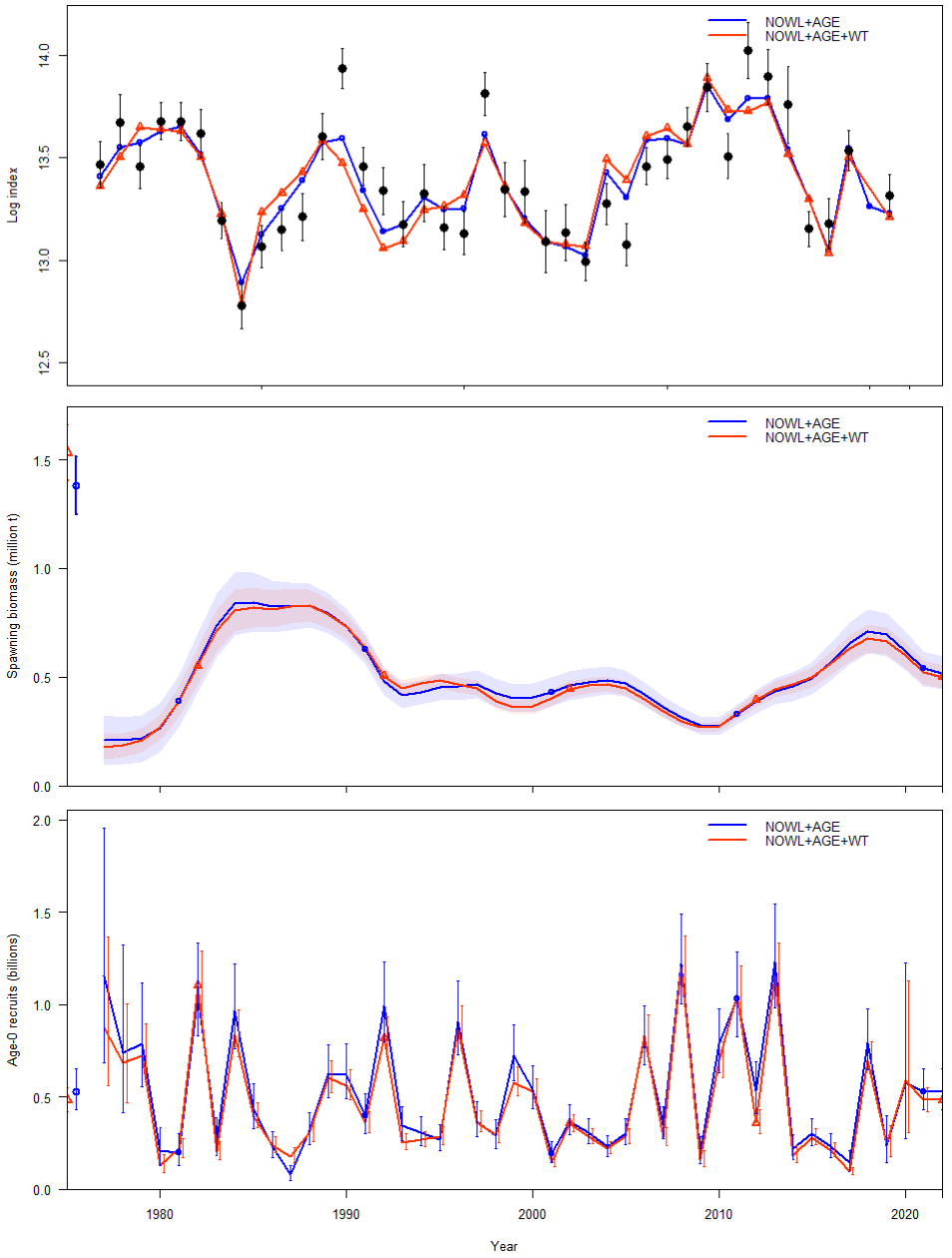




+WT model changes

- Poorer fit to abundance indices
- Composition fits visually indistinguishable
- Minor changes in recruitment and spawning stock biomass estimates

Model 19.12A



Author Recommendation

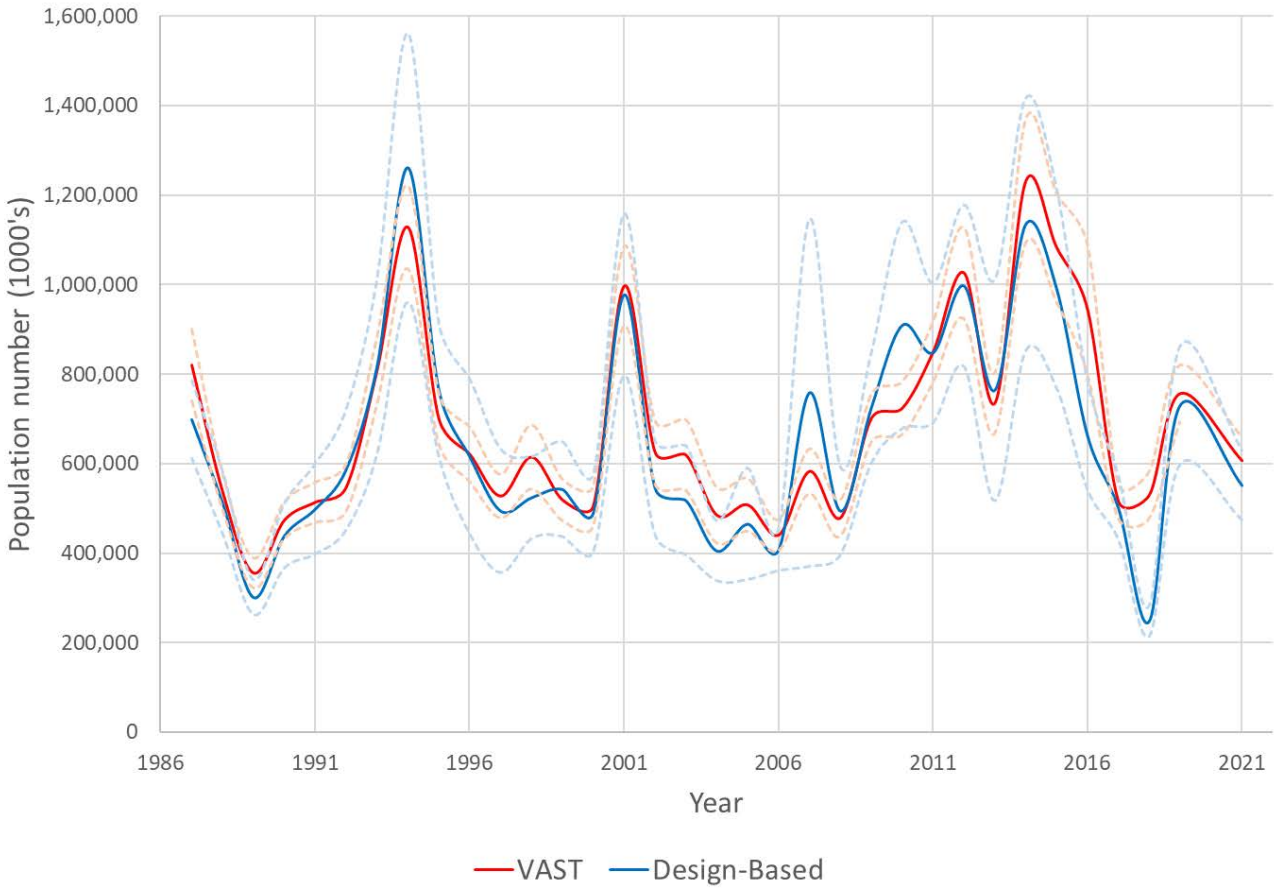
I recommend that the new weighting of the length composition data be considered for 2022, however acceptance of the new weighting be examined more thoroughly once the new 2022 survey and fishery data are added to the model with further examination of model stability and sensitivity to this change.

In addition, I recommend alternative means for calculating the length and age composition input sample sizes be explored in 2023 including bootstrap and VAST derived effective sample sizes.

Fitting additional standard error to abundance indices (+SE)

- VAST bottom trawl survey index variance is ½ of design-based
- VAST survey indices are model output and the index time series can change annually with new data
- Fitting additional variance for indices is commonly used and implemented in Stock Synthesis

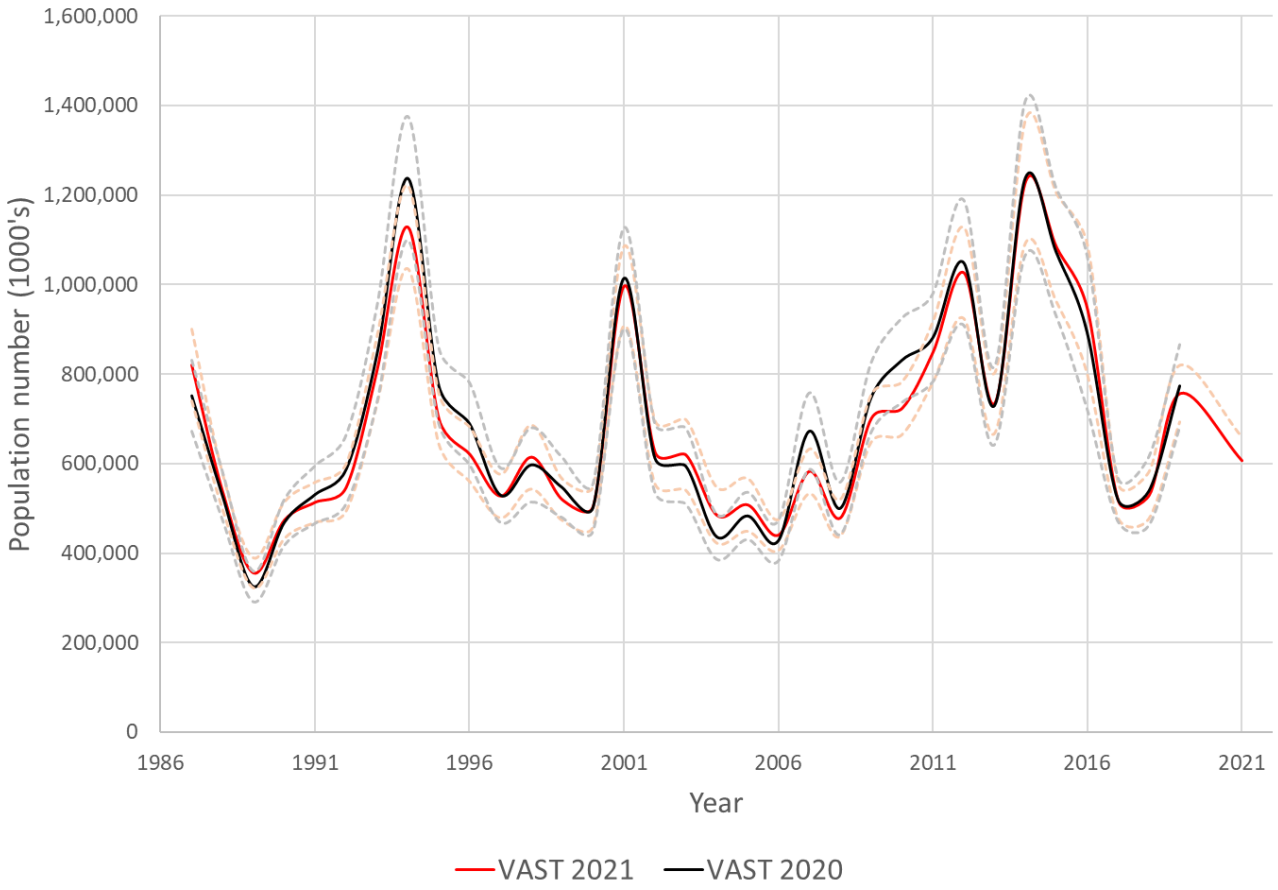
Fitting additional standard error to abundance indices (+SE)



Standard error sets the approximate weight of an index in the model compared to other data sets

VAST standard error is ~1/2 design based

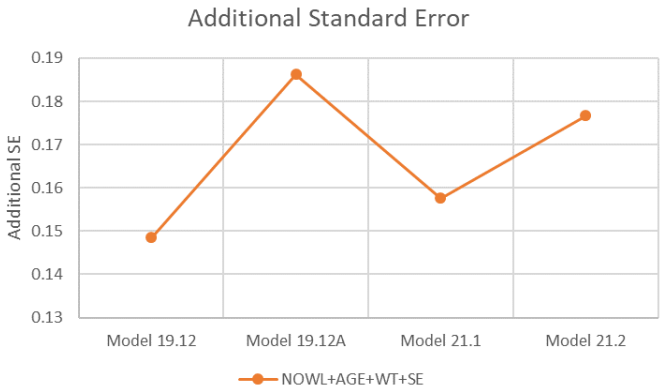
Fitting additional standard error to abundance indices (+SE)



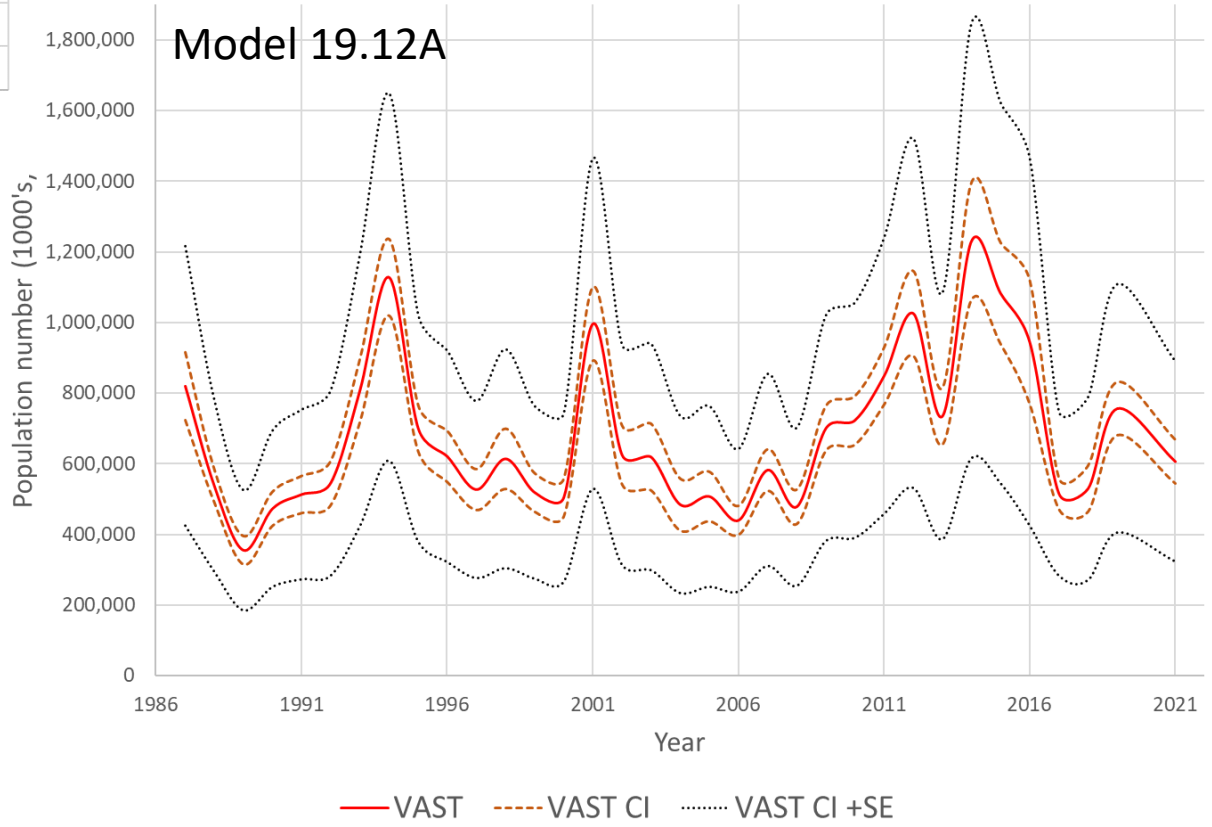
<u>VAST indices are modeled products</u>

Additional data will result in changes to the modeled spatial autocorrelation and therefore changes to the full time series.

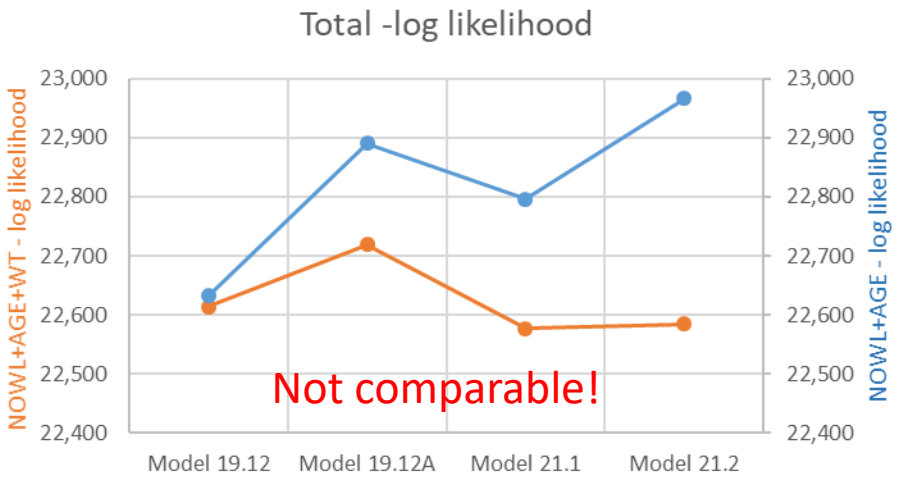
Fitting additional standard error to abundance indices (+SE)

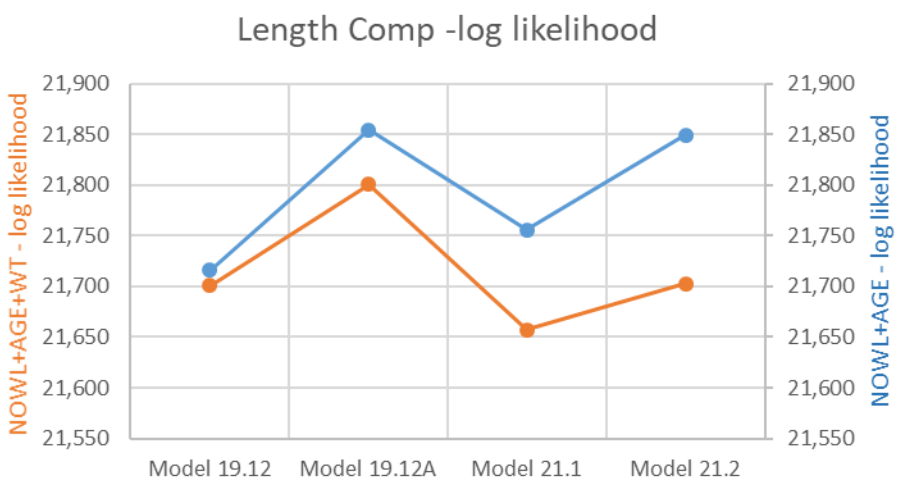


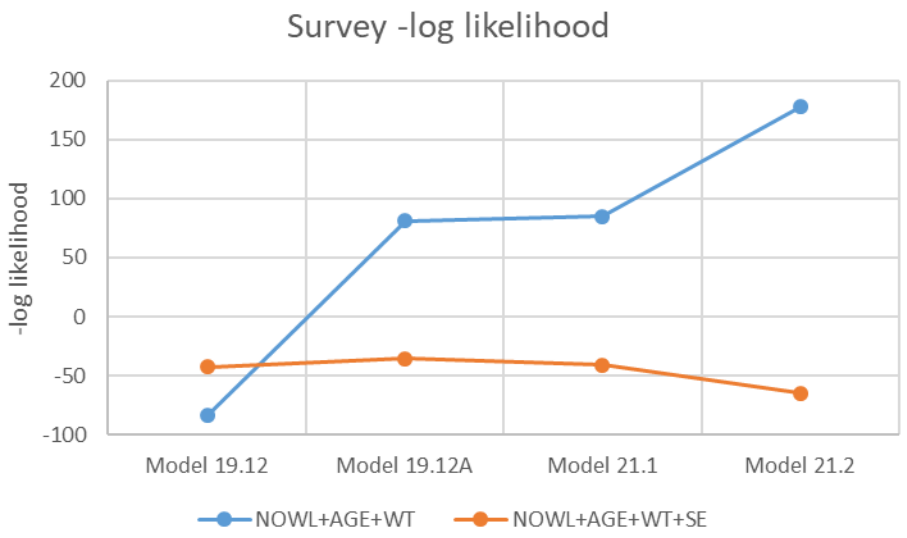
- Inflates variance higher than design-based
- Seems unreasonably high suggesting possible model misspecification

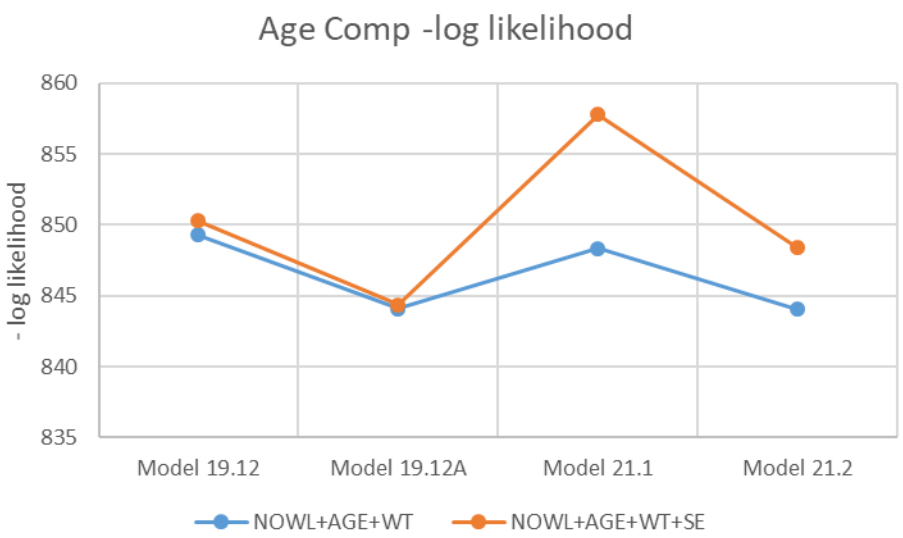


Additional standard error (+SE)



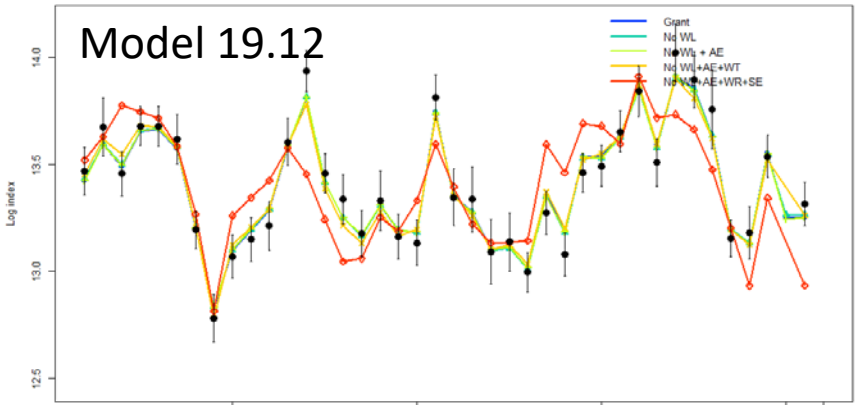


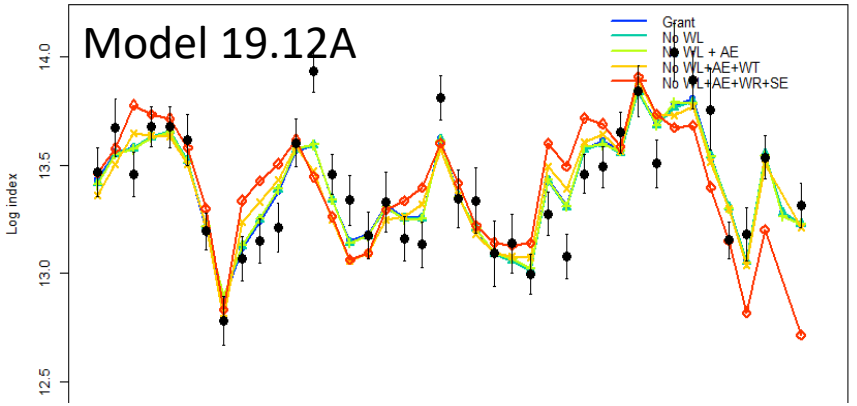


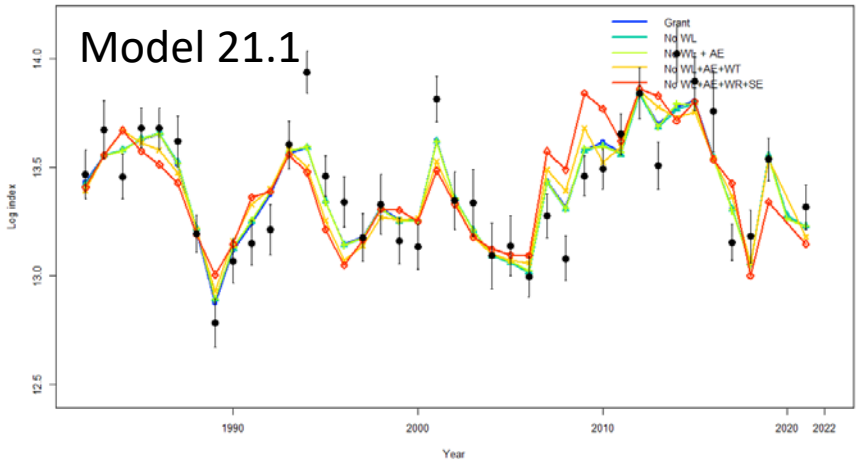


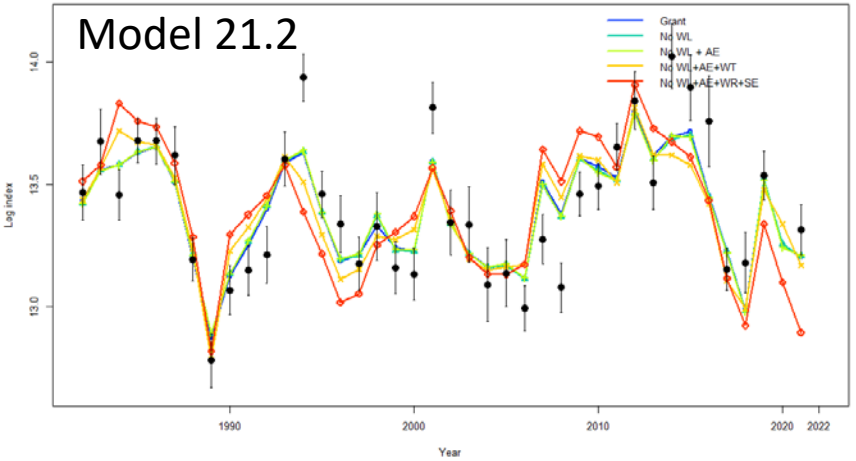
Fitting additional standard error to abundance indices (+SE)

- Substantially degraded fit to the survey index



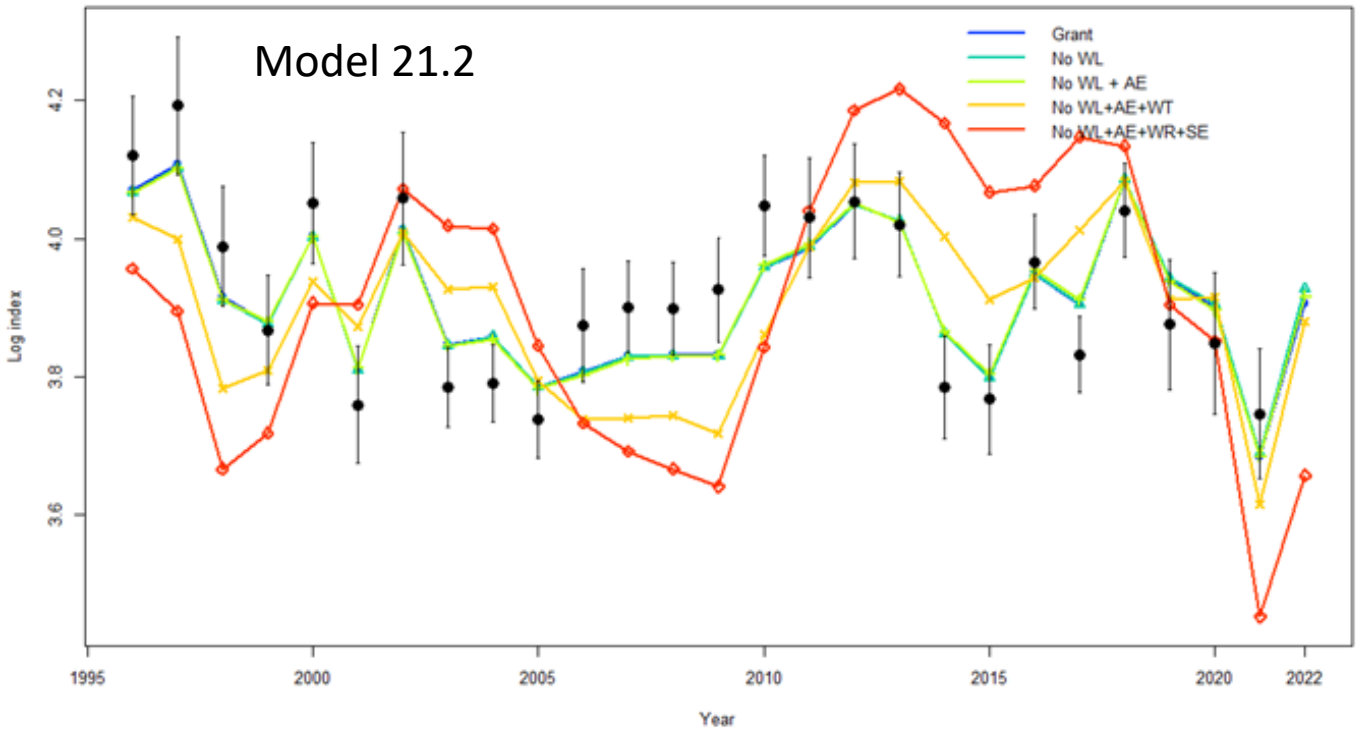




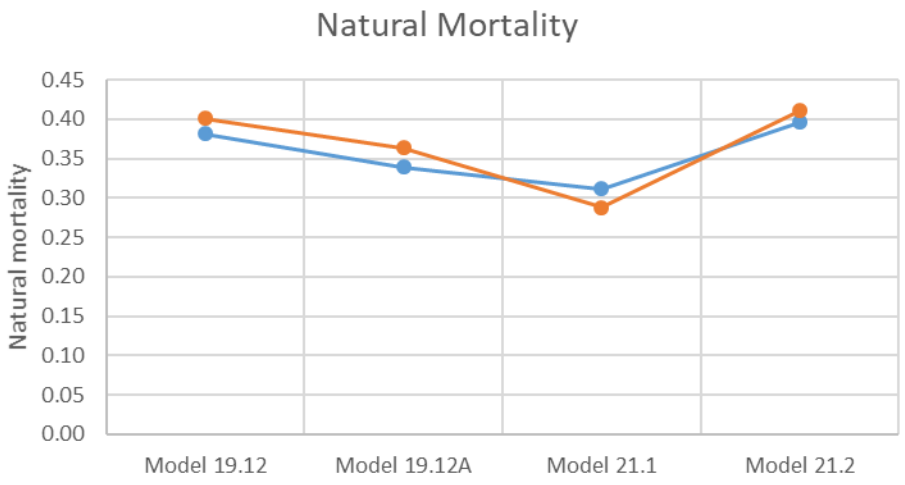


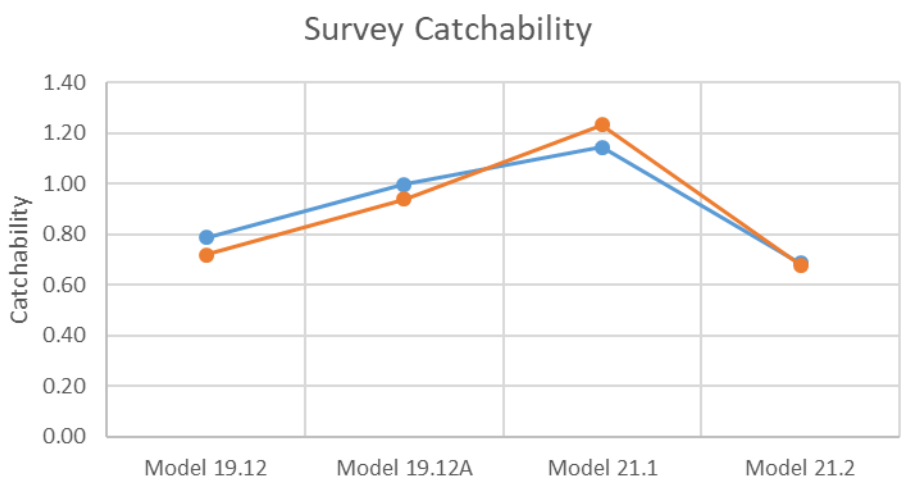
Fitting additional standard error to abundance indices (+SE)

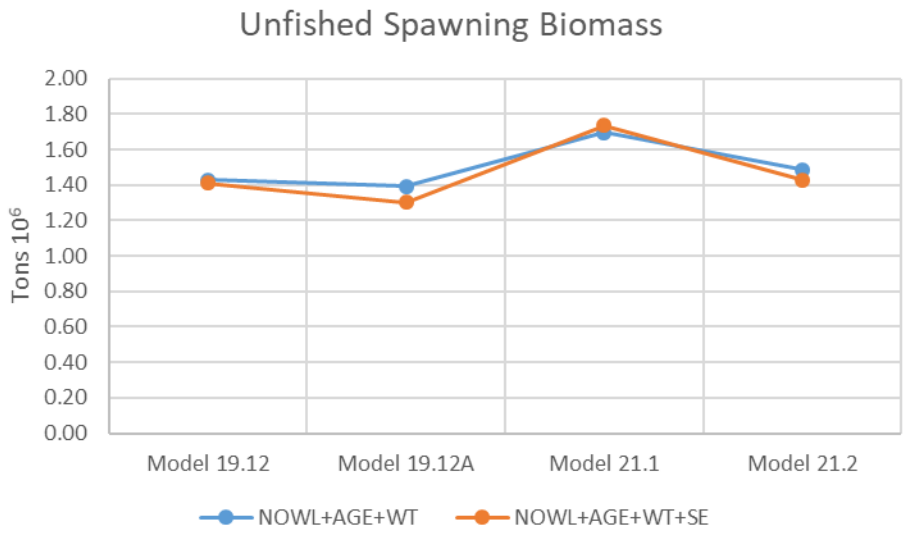
- Substantially degraded fit to the CPUE index for Model 21.2

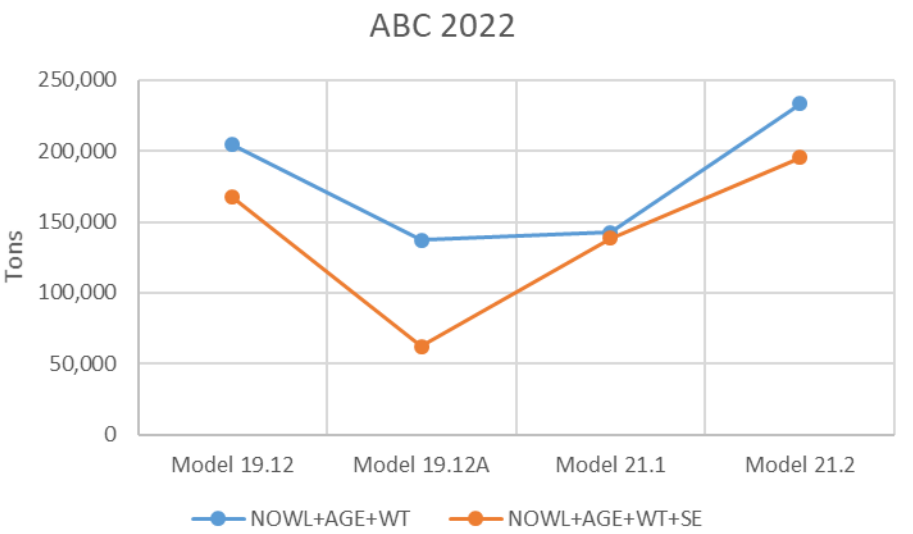


Additional standard error (+SE)

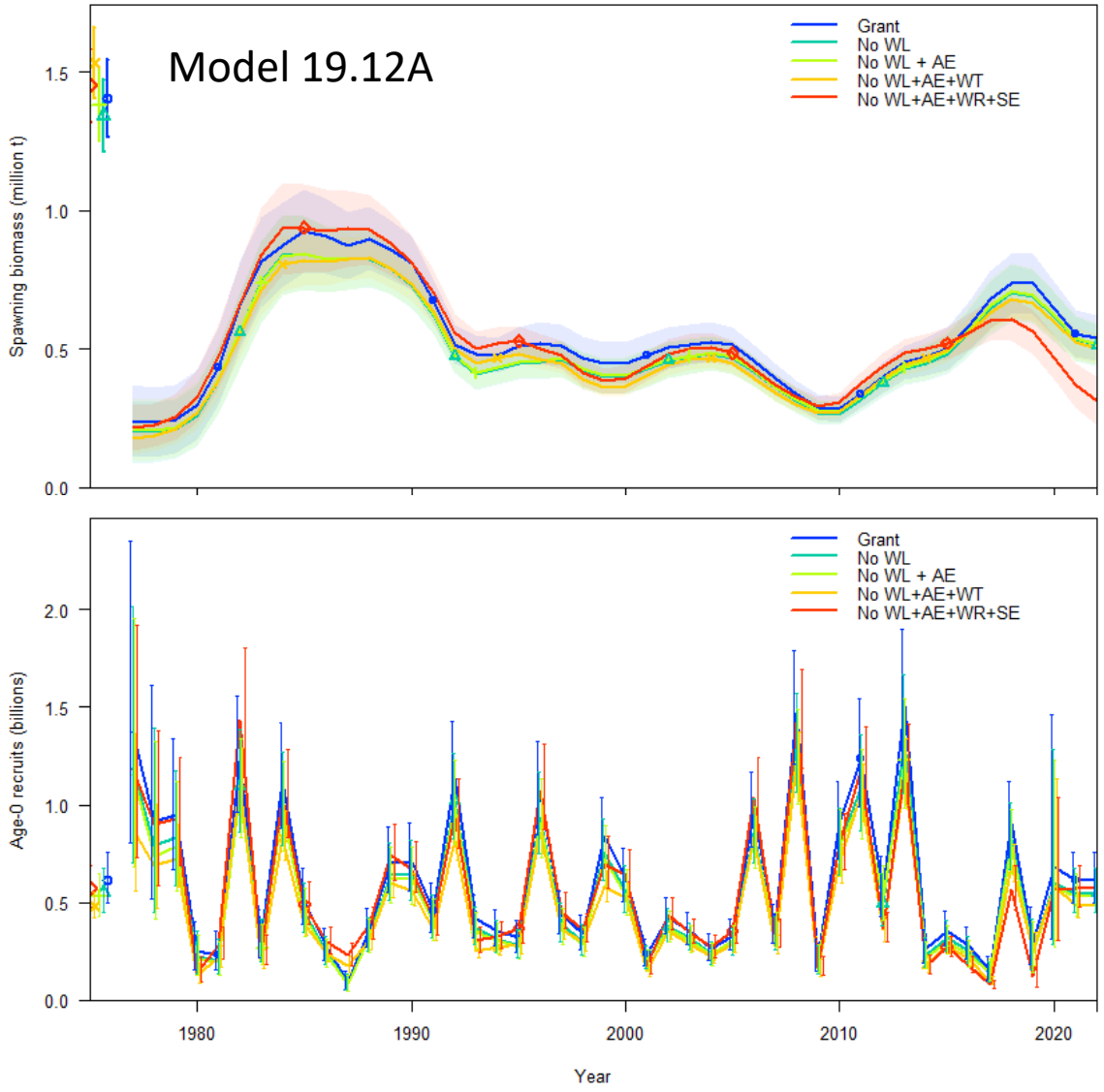








Additional standard error (+SE)

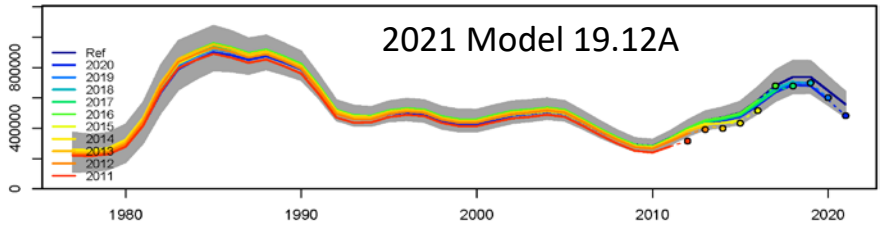


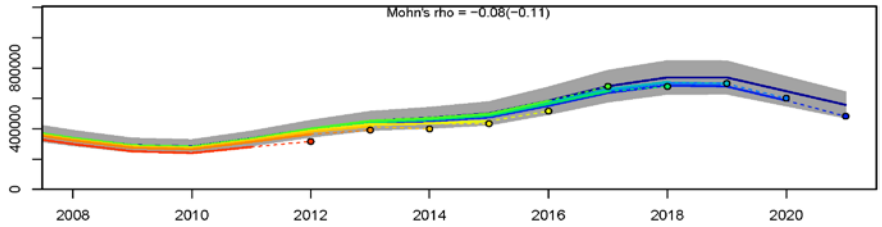
Much lower estimates of recent spawning biomass.

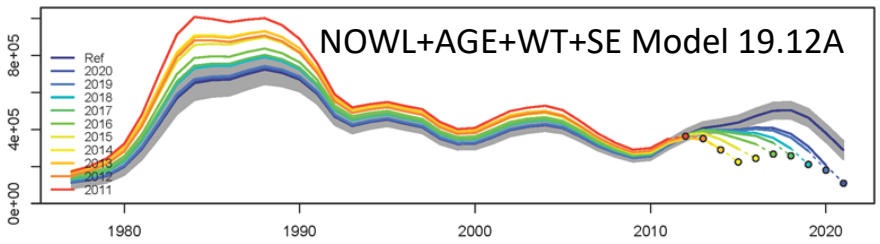
Improved fit to length composition at the cost of poor fit to survey.

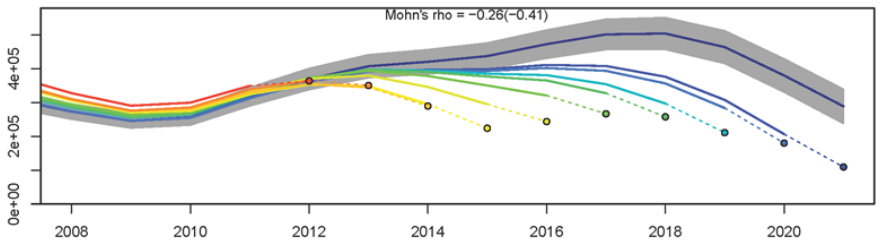
Theory: Change in growth and weight at length leading to model misspecification.

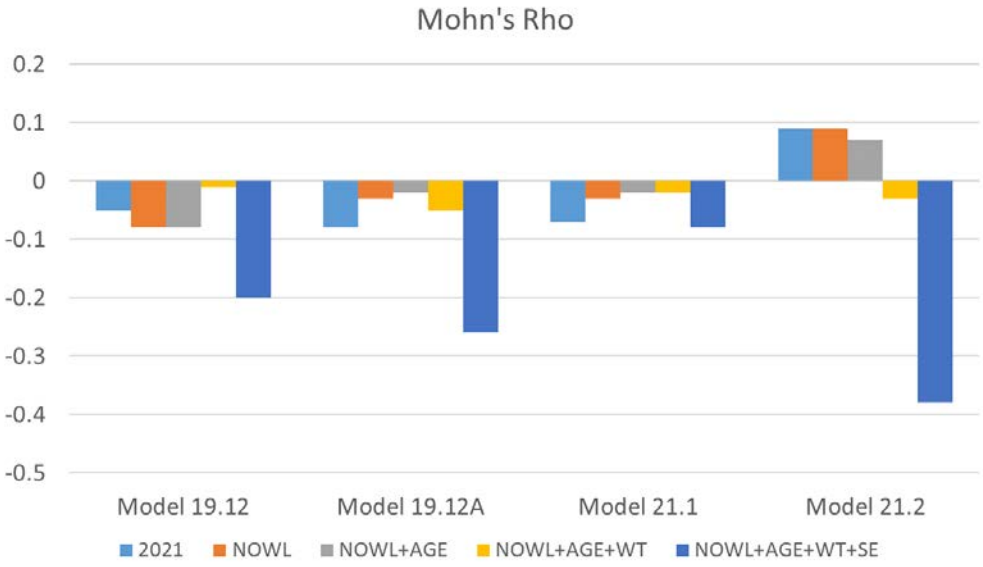
Poorer retrospective pattern with additional standard error (+SE) in all models











Author Recommendation

I recommend that fitting additional standard error to the indices not be adopted for this year's set of ensemble models. Additional exploration of proper variance attribution of VAST indices within the assessment model should continue to be explored in 2023.

Additional observations on current ensemble

- Joint index residual plots produced with and joint mean squared error
- Residual runs tests were performed to examine the distribution of the residuals and whether the residuals were randomly distributed
- The Mean absolute scaled error (MASE) values examine the prediction skill of the models and versions, values greater than 1.0 indicated performance worse than a random walk.

From Carvalho et al. (2021) implemented in the ss3diags R library

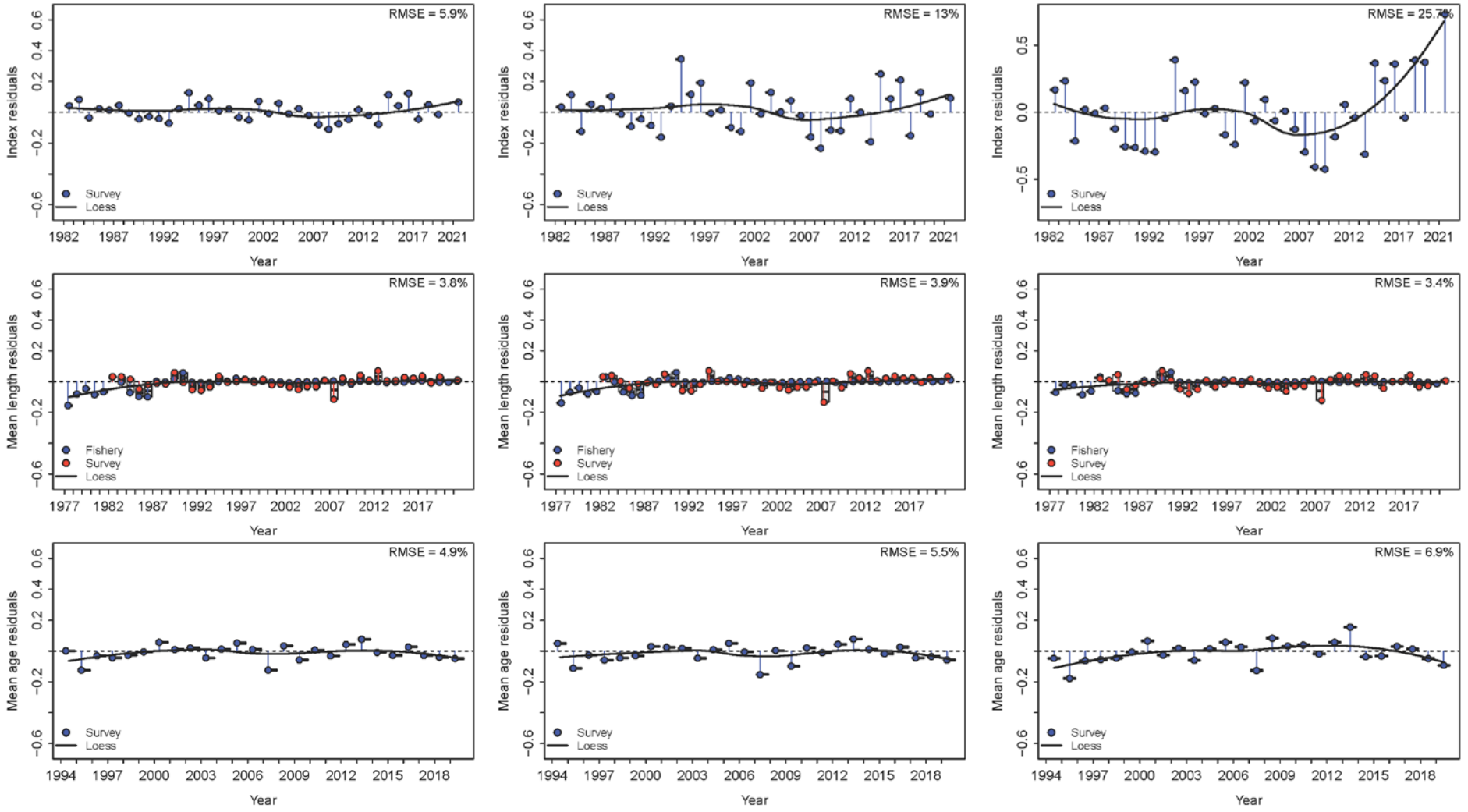
Joint-index residual plots

Model 19.12A

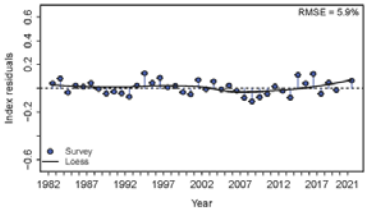
Model 19.12 2021

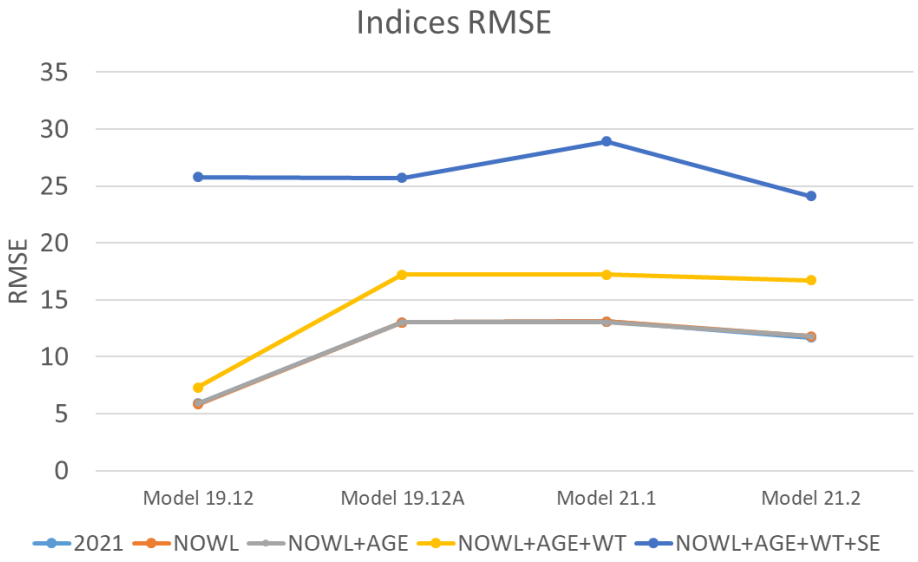
2021

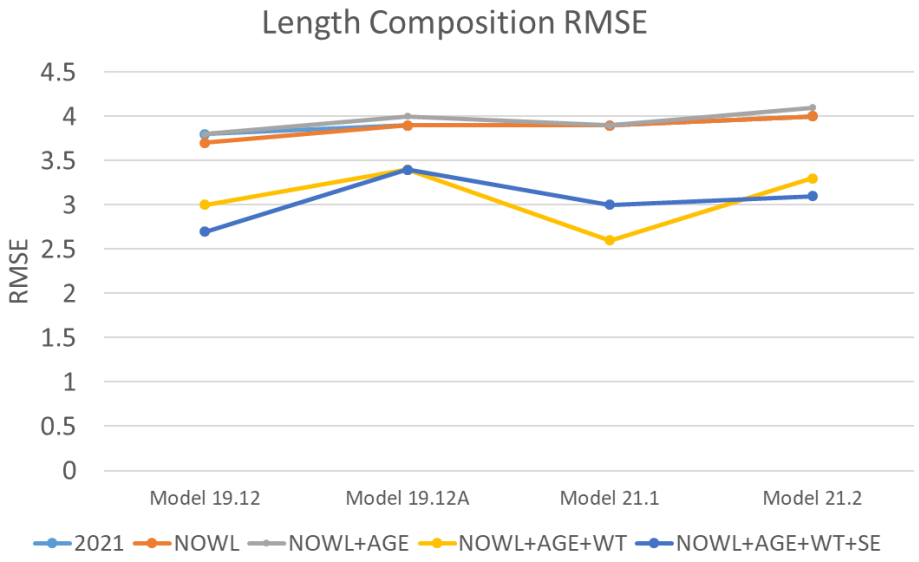
NOWL+AGE+WT+SE

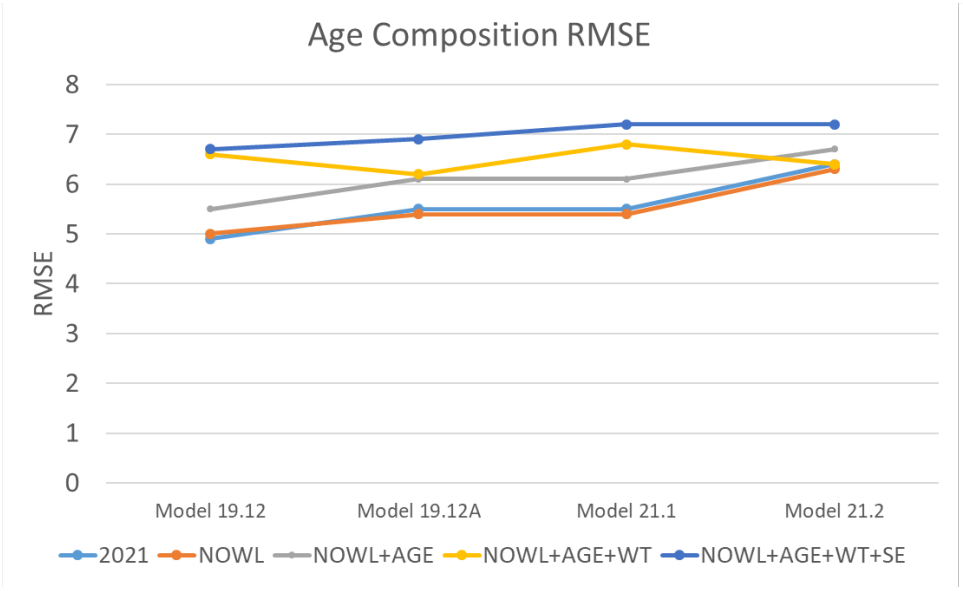


Joint RMSEs


RMSE = 5.9%
Index residuals
0.6
0.4
0.2
-0.2
-0.6
Survey
Loess
1982
1987
1992
1997
2002
2007
2012
2017
2021
Year


Indices RMSE
RMSE
35
30
25
20
15
10
5
0
Model 19.12
Model 19.12A
Model 21.1
Model 21.2
2021
NOWL
NOWL+AGE
NOWL+AGE+WT
NOWL+AGE+WT+SE

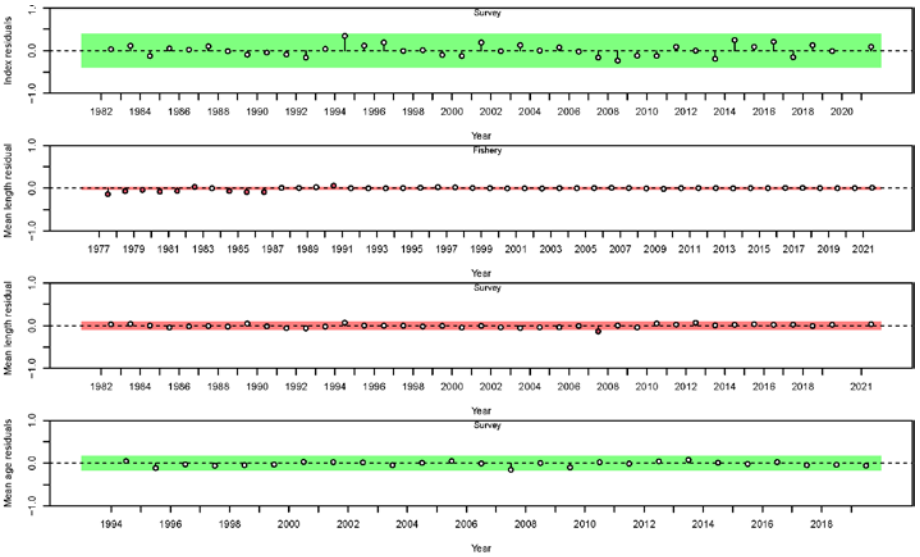

Length Composition RMSE
RMSE
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
Model 19.12
Model 19.12A
Model 21.1
Model 21.2
2021
NOWL
NOWL+AGE
NOWL+AGE+WT
NOWL+AGE+WT+SE


Age Composition RMSE
RMSE
8
7
6
5
4
3
2
1
0
Model 19.12
Model 19.12A
Model 21.1
Model 21.2
2021
NOWL
NOWL+AGE
NOWL+AGE+WT
NOWL+AGE+WT+SE

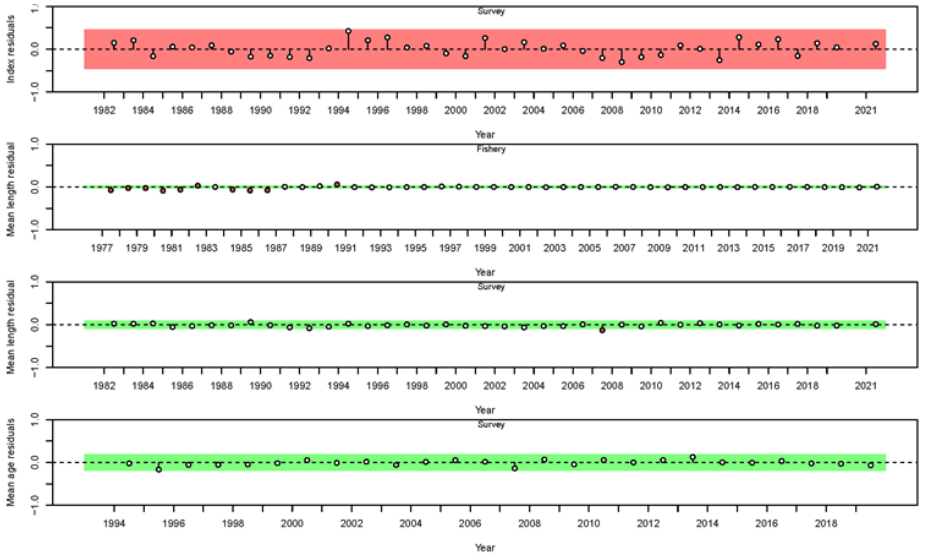
Residual runs tests

Model 19.12A

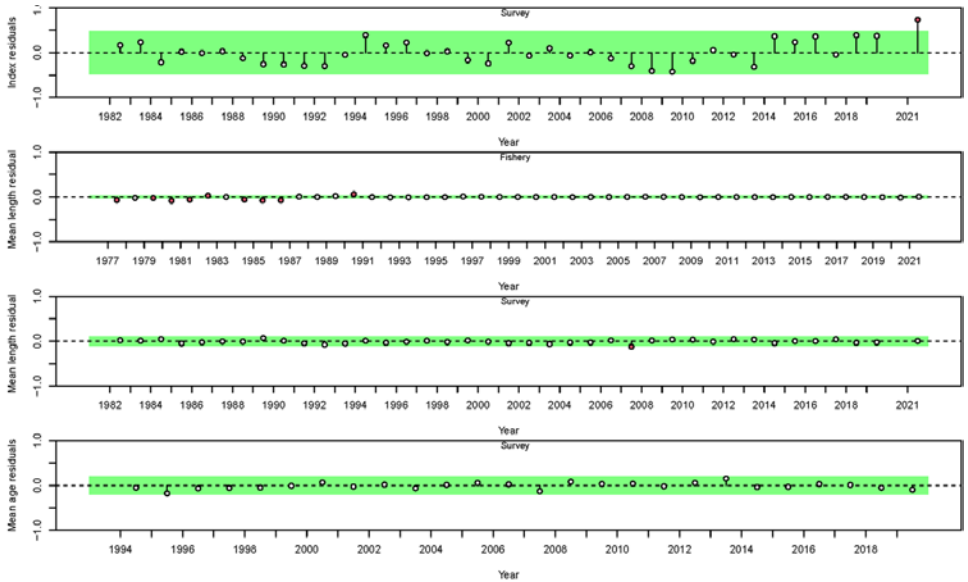
2021 and NOWL and NOWL+AGE



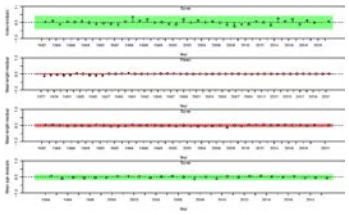
NOWL+AGE+WT



NOWL+AGE+WT+SE

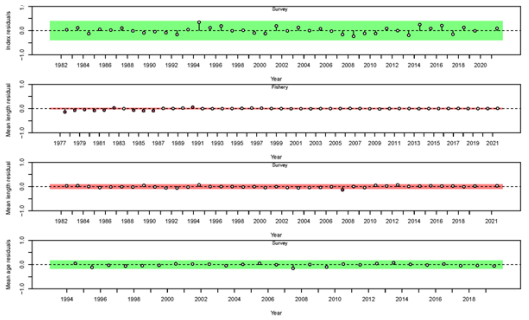


Residual run tests



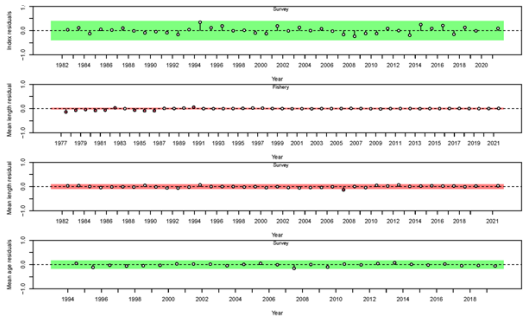
- Mixed results across models and versions for the index and length composition tests
- All models and versions passed the age composition tests
- By version across all models (17 tests each version)
  - NOWL+AGE+WT+SE performed the best with 4 failures
  - NOWL+AGE+WT was next with 5 failures
  - The remaining versions had 8 failures each, but no consistency in which data components
- By Model across all versions (20 tests each)
  - All models except 21.2 performed equally with 7 failures
  - Model 21.2 had 10 total failures +2 for the fishery CPUE index

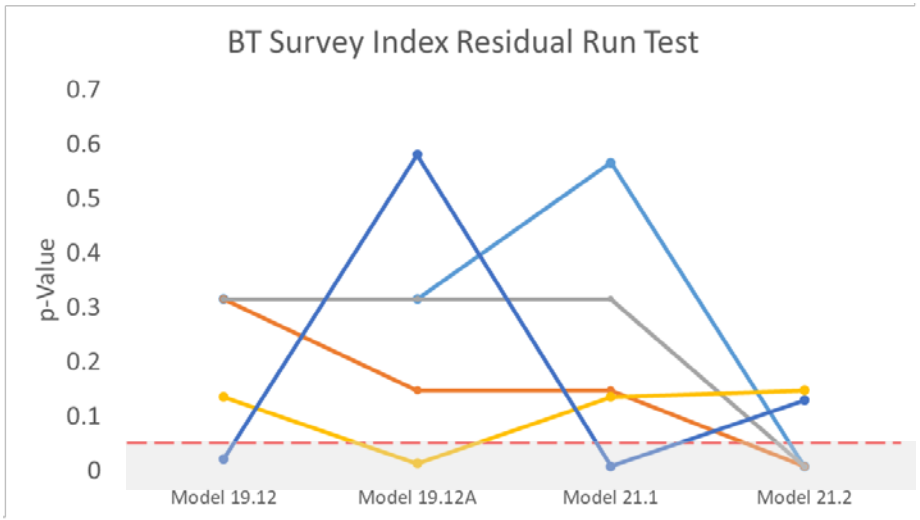
Residual runs test: Fishery mean Length

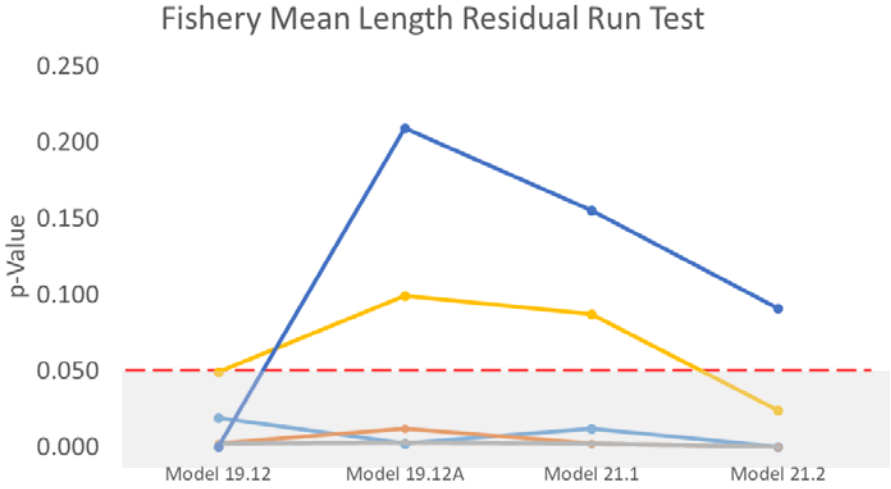


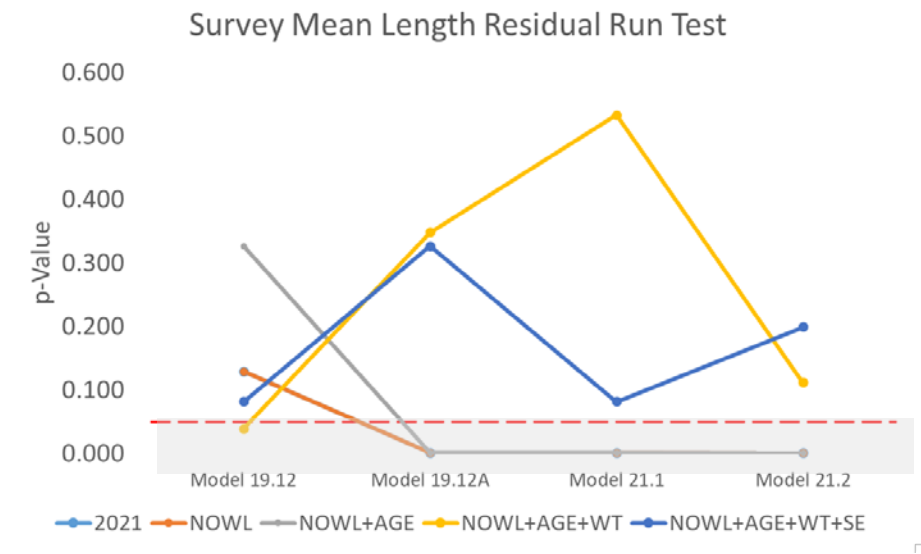
Version	*Model 19.12*	*Model 19.12A*	*Model 21.1*	*Model 21.2*	*Label*
2021	0.315	0.315	0.566	0.008	BT Survey index
NOWL	0.315	0.147	0.147	0.008	BT Survey index
NOWL+AGE	0.315	0.315	0.315	0.008	BT Survey index
NOWL+AGE+WT	0.135	0.013	0.135	0.147	BT Survey index
NOWL+AGE+WT+SE	0.021	0.58	0.008	0.129	BT Survey index
2021				0.120	Fishery Index
NOWL				0.120	Fishery Index
NOWL+AGE				0.120	Fishery Index
NOWL+AGE+WT				0.024	Fishery Index
NOWL+AGE+WT+SE				0.000	Fishery Index
2021	0.019	0.002	0.012	0.000	Fishery Length
NOWL	0.002	0.012	0.002	0.000	Fishery Length
NOWL+AGE	0.002	0.003	0.002	0.000	Fishery Length
NOWL+AGE+WT	0.049	0.099	0.087	0.024	Fishery Length
NOWL+AGE+WT+SE	0.000	0.209	0.155	0.091	Fishery Length
2021	0.129	0.001	0.001	0.000	Survey Length
NOWL	0.129	0.001	0.001	0.000	Survey Length
NOWL+AGE	0.326	0.001	0.001	0.000	Survey Length
NOWL+AGE+WT	0.039	0.348	0.533	0.111	Survey Length
NOWL+AGE+WT+SE	0.081	0.326	0.081	0.199	Survey Length
2021	0.512	0.512	0.512	0.08	Survey Age
NOWL	0.512	0.512	0.512	0.08	Survey Age
NOWL+AGE	0.704	0.057	0.057	0.219	Survey Age
NOWL+AGE+WT	0.355	0.355	0.448	0.541	Survey Age
NOWL+AGE+WT+SE	0.704	0.355	0.355	0.355	Survey Age

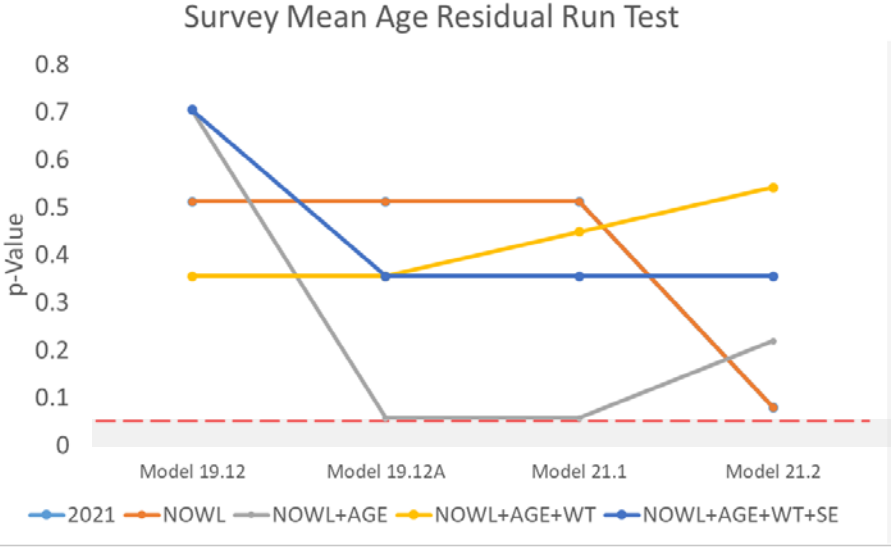
Residual runs test: Fishery mean Length





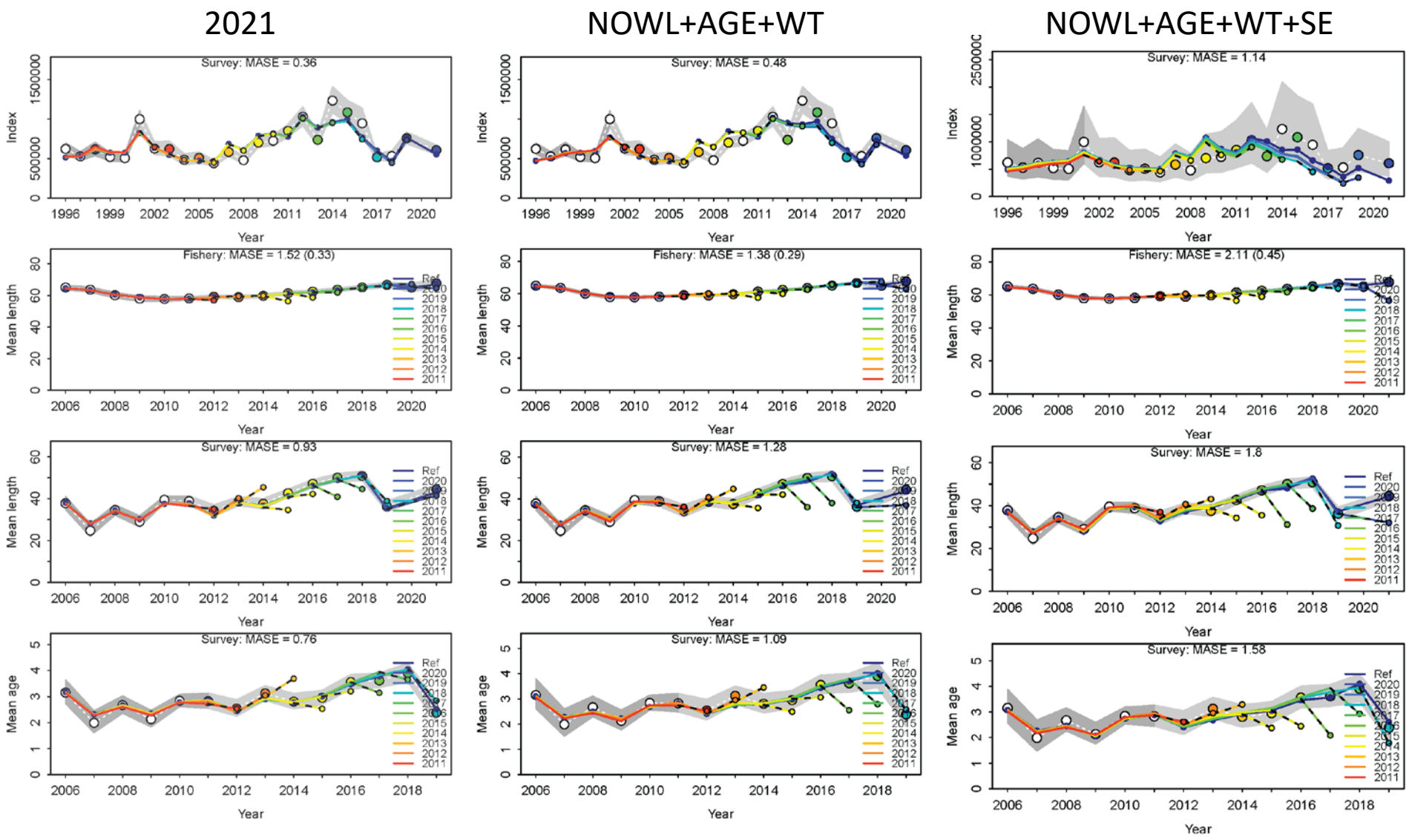






Mean absolute scaled error (MASE)

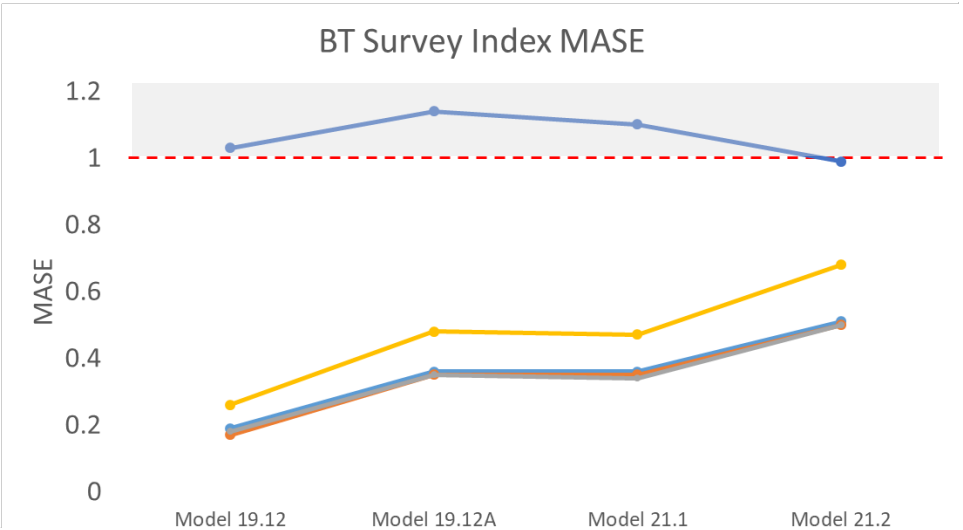
Model 19.12A

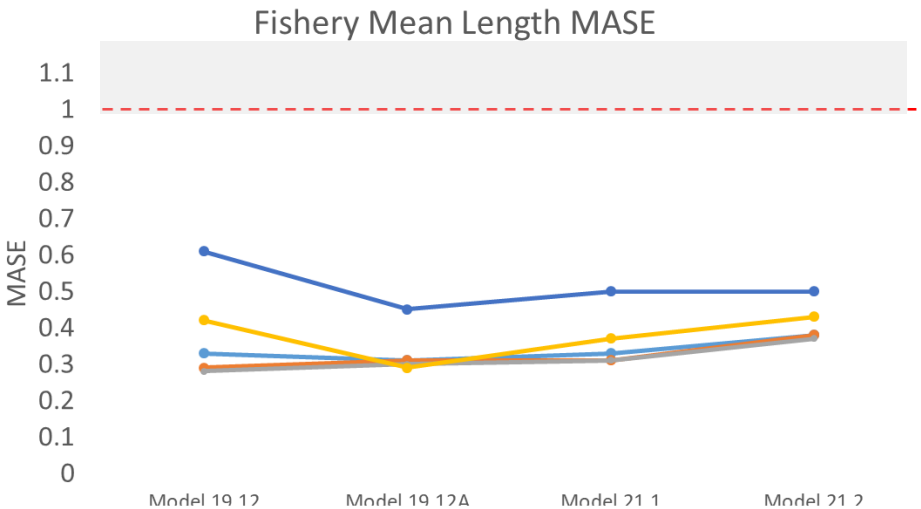


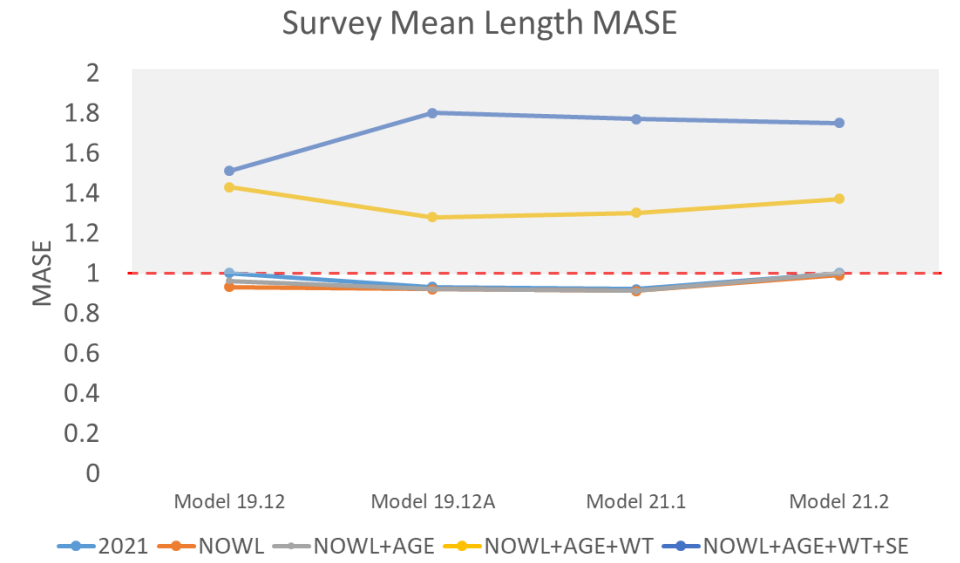
Mean absolute scaled error (MASE)

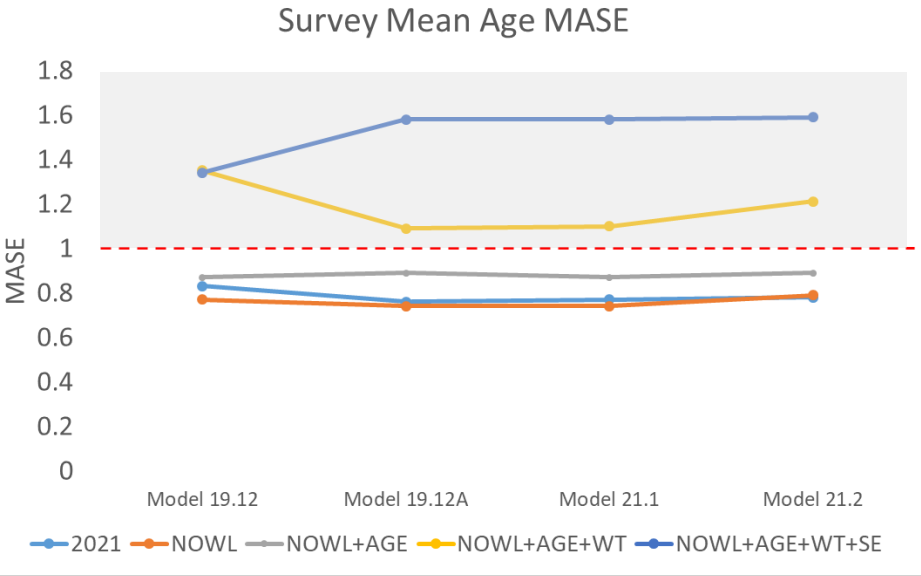
Version	Model 19.12	Model 19.12A	Model 21.1	Model 21.2	Label
2021	0.19	0.36	0.36	0.51	BT Survey Index
NOWL	0.17	0.35	0.35	0.50	BT Survey Index
NOWL+AGE	0.18	0.35	0.34	0.50	BT Survey Index
NOWL+AGE+WT	0.26	0.48	0.47	0.68	BT Survey Index
NOWL+AGE+WT+SE	1.03	1.14	1.10	0.99	BT Survey Index
2021				0.55	CPUE Index
NOWL				0.53	CPUE Index
NOWL+AGE				0.47	CPUE Index
NOWL+AGE+WT				1.04	CPUE Index
NOWL+AGE+WT+SE				2.46	CPUE Index
2021	0.33	0.31	0.33	0.38	Fishery Mean Length
NOWL	0.29	0.31	0.31	0.38	Fishery Mean Length
NOWL+AGE	0.28	0.30	0.31	0.37	Fishery Mean Length
NOWL+AGE+WT	0.42	0.29	0.37	0.43	Fishery Mean Length
NOWL+AGE+WT+SE	0.61	0.45	0.50	0.50	Fishery Mean Length
2021	1.00	0.93	0.92	1.00	Survey Mean Length
NOWL	0.93	0.92	0.91	0.99	Survey Mean Length
NOWL+AGE	0.96	0.92	0.91	1.00	Survey Mean Length
NOWL+AGE+WT	1.43	1.28	1.30	1.37	Survey Mean Length
NOWL+AGE+WT+SE	1.51	1.80	1.77	1.75	Survey Mean Length
2021	0.83	0.76	0.77	0.78	Survey Mean Age
NOWL	0.77	0.74	0.74	0.79	Survey Mean Age
NOWL+AGE	0.87	0.89	0.87	0.89	Survey Mean Age
NOWL+AGE+WT	1.35	1.09	1.10	1.21	Survey Mean Age
NOWL+AGE+WT+SE	1.34	1.58	1.58	1.59	Survey Mean Age

Mean absolute scaled error (MASE)









Author Recommendation

I recommend that the authors in 2023 re-explore a seasonal model for Bering Sea Pacific cod and in light of the most recent genetic and tagging data (McDermott personal comm.) explore an expanded spatial model that incorporates the western Gulf of Alaska in the model.