NOAA FISHERIES

Alaska Fisheries Science Center

Preliminary assessment of Pacific cod in the Eastern Bering Sea

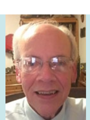
Grant Thompson, Steve Barbeaux, Jason Conner, Cecilia O’Leary, Ingrid Spies, and James Thorson

September 22, 2021

CIE review

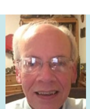
Overview

- Dates: April 26-30, 2021
- Reviewers:
  - Yan Jiao, Virginia Polytechnic Institute and State University
  - Arni Magnusson, General Fisheries Commission of the Mediterranean (FAO)
  - Henrik Sparholt, University of Copenhagen
- Chair: Ingrid Spies
- Assessment team: Grant Thompson and Steven Barbeaux
- Original terms of reference, plan for conduct of the meeting, background documents, and full reports of the reviewers: https://apps-afsc.fisheries.noaa.gov/Plan_Team/2021_pcod_cie/
- Attachment 2.1.1 summarizes 50 reviewer comments, with responses



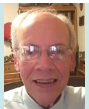
Terms of reference (1 of 2)

- Terms of reference consisted of six main topics
- Each topic included 3 subtopics consisting of recommendations from:
  - Groundfish Plan Team
  - Scientific and Statistical Committee
  - Alistair Dunn (FLC consultant)
- The reviewers added a fourth subtopic of their own to the “Other” topic
- As there was insufficient time to address all topics and subtopics, the reviewers were asked to prioritize them
- This resulted in some subtopics receiving no recommendations from the reviewers, as expected



Terms of reference (2 of 2)

- As prioritized by the reviewers, the six topics were:
  1. Ensemble modeling
  2. Movement
  3. Fishery CPUE
  4. Age data
  5. Compositional data
  6. Other
- Most of the discussion focused on “Ensemble Modeling,” especially:
  - Development of the specific models to include in the ensemble
  - Specification of model weights



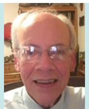
The CIE ensemble

- Reviewers adopted a “+1” approach to specify the set of models
  - Base model = Model 19.12a
  - Four new features
  - Each new model = base model + 1 new feature
  - Four new features $\rightarrow$ 5 models overall
    - Factorial design would have yielded 2^4=16 models overall

Feature	19.12a	19.12	"20.8a"	"20.9a"	"21.cie"
Feature 1: Allow catchability to vary?	no	yes	no	no	no
Feature 2: Allow domed survey selectivity?	no	no	yes	no	no
Feature 3: Use fishery CPUE?	no	no	no	yes	no
Feature 4: Estimate survey CV internally?	no	no	no	no	yes
Model weight:	0.2459	0.2213	0.1803	0.1311	0.2213

- The above will be covered in detail later in the presentation

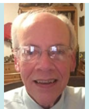
NOAA FISHERIES



Team and SSC comments

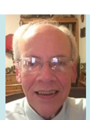
Comments on assessments in general (1 of 2)

- “Risk table” recommendations will be addressed in the final draft
- SSC1: “The SSC cautions against standardized model fitting (e.g., a single error distribution, set of covariates, number of knots), other than as a starting point…. It is more important for each species to have a statistically rigorous model selection process resulting in good model fit and diagnostics than the simplicity of fitting the same approach to all species: unlike design-based estimators, the SSC suggests that one size does not fit all for VAST models. For each species, assessment documents should describe why the particular error distributions, covariates, and number of knots were chosen for that individual species.”
    - *Response:* An evaluation of alternative configurations for the VAST model of trawl survey index data is presented under “VAST estimates of survey abundance” in the “Data” section.



Comments on assessments in general (2 of 2)

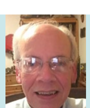
- SSC2: "In general, …the SSC recommends the continued inclusion of community engagement and dependency indices at varying scales in ESPs, ESRs, and SAFEs. For ESPs specifically, changes in patterns of community engagement and dependency at the stock level have the potential to inform not only stock assessments and analyses that support fishery management, but they may also function as early indicators of larger ecosystem changes."
  - *Response:* See response listed in the revised ESP



Comments specific to this assessment (1 of 8)

- GPT1: "The Team recommended that the fishery CPUE be standardized using alternative statistical methods and that it be discussed at the CIE review in 2021. This should also include a discussion of historical changes in the fishery that may affect the relationship of the index to abundance."
  - *Response:* A first attempt at standardizing fishery CPUE using alternative statistical methods has been completed (see "VAST estimates of fishery catch per unit effort" in the "Data" section). This comment was forwarded to the CIE reviewers for their consideration (Attachment 2.1.1). In response, one result of the CIE review was the inclusion in this preliminary assessment of a model incorporating the new index of fishery CPUE (see "Alternative models" in the "Models" section).

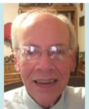
NOAA FISHERIES



Comments specific to this assessment (2 of 8)

- GPT2: “The Team recommended collating fishery information in the ESP. Although the CPUE index was of concern to the Team, the Team recognizes that fishery performance has been improving and that these observations should not be ignored. Inclusion of fishery performance in the ESP and evaluation of the CPUE index with those performance metrics may help provide important insights.”
  - *Response:* Collating *all* fishery information in the ESP could prove awkward, because some of it is routinely used by the assessment models (e.g., catch and fishery size composition), and it seems more appropriate to collate such data in the main text of the assessment. Moreover, because fishery CPUE data are used in one of the assessment models, it seems more efficient to collate all fishery CPUE information in the main text than to split this information into two parts, with one part placed in the main text and the other in the ESP. The revised ESP continues to include aspects of fishery performance other than fishery CPUE.

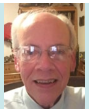
NOAA FISHERIES



Comments specific to this assessment (3 of 8)

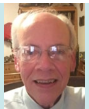
- GPT3: “The Team recommended the following topics could be considered for the 2021 CIE review: development of a fishery CPUE index, incorporation of dome-shaped survey selectivity, models to include in an ensemble, whether to apply the sloping HCR before or after ensemble averaging of SSB and other reference points, and development of movement models.”
    - *Response:* These topics were considered during the CIE review. In response, some results of the CIE review were the inclusion in this preliminary assessment of: 1) a model incorporating a new index of fishery CPUE (see response to comment GPT1), 2) a model incorporating dome-shaped survey selectivity, 3) a set of models to include in an ensemble, 4) a simple conceptual model (*not* part of the ensemble) addressing movement of Pacific cod between American and Russian jurisdictions (Attachment 2.1.2), and 5) further analysis of whether to apply the harvest control rule before or after model averaging (Attachment 2.1.3)

NOAA FISHERIES



Comments specific to this assessment (4 of 8)

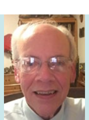
- SSC3: “The SSC supports items proposed by the BSAI GPT for inclusion in the CIE review of this assessment planned for 2021. Proposed topics include: development of a standardized fishery CPUE index using alternative statistical methods, incorporation of dome-shaped survey selectivity, discussion of models to include in an ensemble, whether to apply the sloping harvest control rule before or after ensemble averaging of SSB and other reference points, and development of movement models.”
    - *Response:* See response to comment GPT3.



Comments specific to this assessment (5 of 8)

- SSC4: "The SSC also recommends consideration of suggestions offered by Alistair Dunn (public comment) about other factors that could be included in the CIE review if time is available including: inclusion of other survey information (e.g., the IPHC and sablefish surveys), and considerations about how best to include the fishery age and size composition data. Additionally, Mr. Dunn suggested that the analysis of fishery CPUE data suggested by the GPT could include development of spatiotemporal analyses of fleet-specific CPUE indices that may help inform the assessment.
  - *Response:* These topics were considered during the CIE review. Although most of them did not result in very many specific recommendations from the reviewers, it may be noted that the new fishery CPUE index mentioned in response to comment GPT1 does involve a spatiotemporal, fleet-specific analysis.

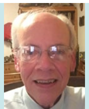
NOAA FISHERIES



Comments specific to this assessment (6 of 8)

- SSC5: “The SSC also encourages review of further efforts to include fishery age data in future analyses.”
  - *Response:* This comment was forwarded for consideration during the CIE review. Although some specific recommendations were received (Attachment 2.1.1), time was insufficient to implement them in this preliminary assessment.
- SSC6: “If time allows, the CIE could comment on avenues for incorporating spatial dynamics and movement.”
  - *Response:* This comment was forwarded for consideration during the CIE review. In response, one result of the CIE review was the inclusion in this preliminary assessment of a conceptual model (*not* part of the ensemble) addressing movement of Pacific cod between American and Russian jurisdictions (Attachment 2.1.2).

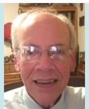
NOAA FISHERIES



Comments specific to this assessment (7 of 8)

- SSC7: “In addition, the SSC would like the CIE review to include an evaluation of the use of ensemble modeling in the NPFMC management system, and specifically whether the structural uncertainty and historical challenges in identifying a robust base model make Pacific cod a good application for ensemble modeling. The SSC acknowledges the trade-off between review capacity and the addition of models comprising an ensemble, but also recognizes that the goals of developing an ensemble that describes a range of structural uncertainties differs from those of refining a single best model.”
  - *Response:* This comment was forwarded for consideration during the CIE review. The reviewers were unanimous in their conclusion that the EBS Pacific cod assessment is a good candidate for ensemble modeling, in part *because* of the structural uncertainty.

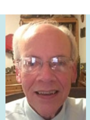
NOAA FISHERIES



Comments specific to this assessment (8 of 8)

- SSC8: “For community harvest revenue indicators, the SSC recommends that the analysts consider aggregating small communities that cannot be individually disclosed into a single indicator that can be displayed along with the limited number of larger community indicators that can be disclosed, for consistency with other ESPs and for the sake of a more comprehensive portrayal of EBS Pacific cod community engagement trends.”
  - *Response:* See response listed in the revised ESP.

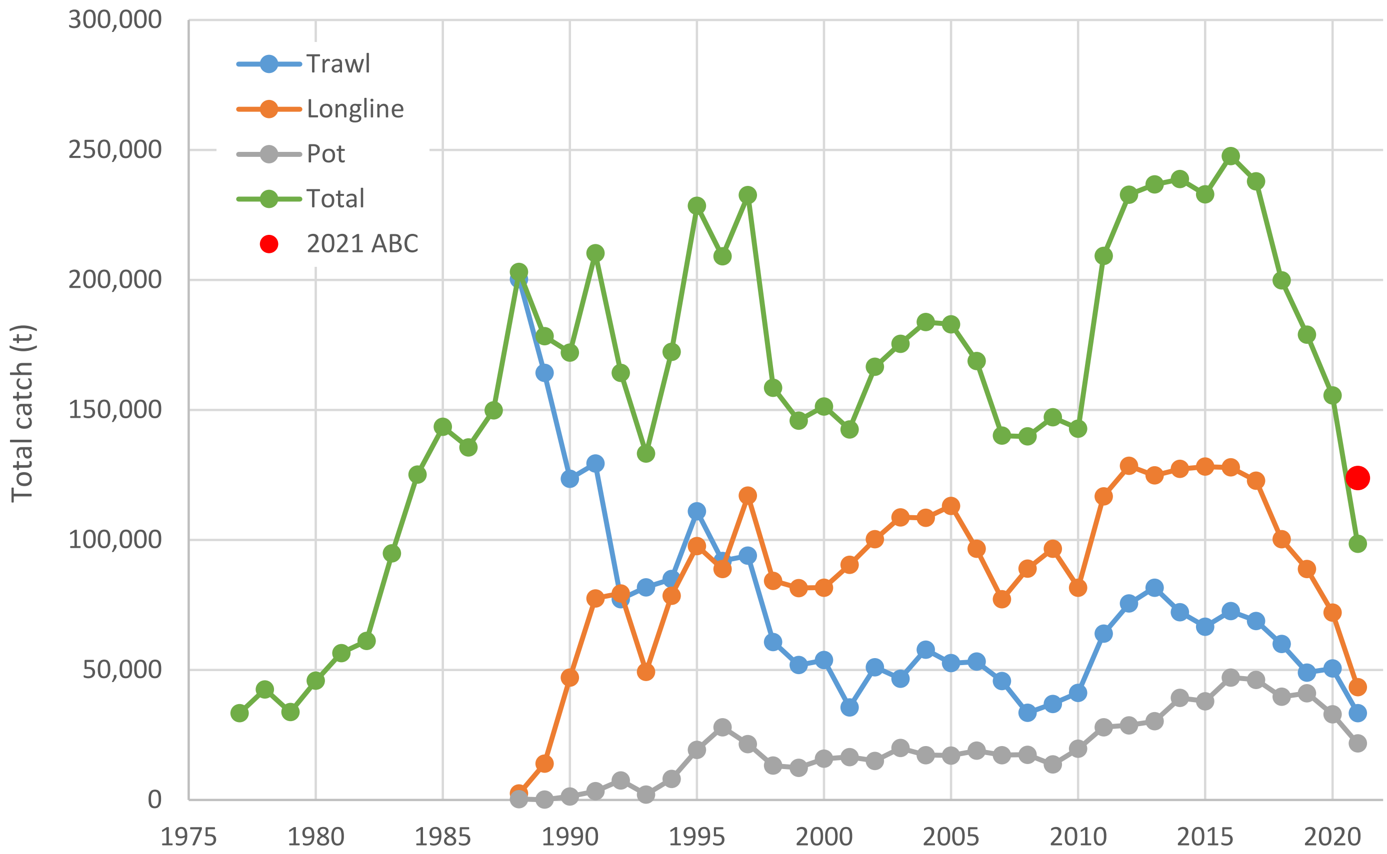
NOAA FISHERIES

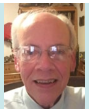


Data

Data for context only: catch (1 of 2)

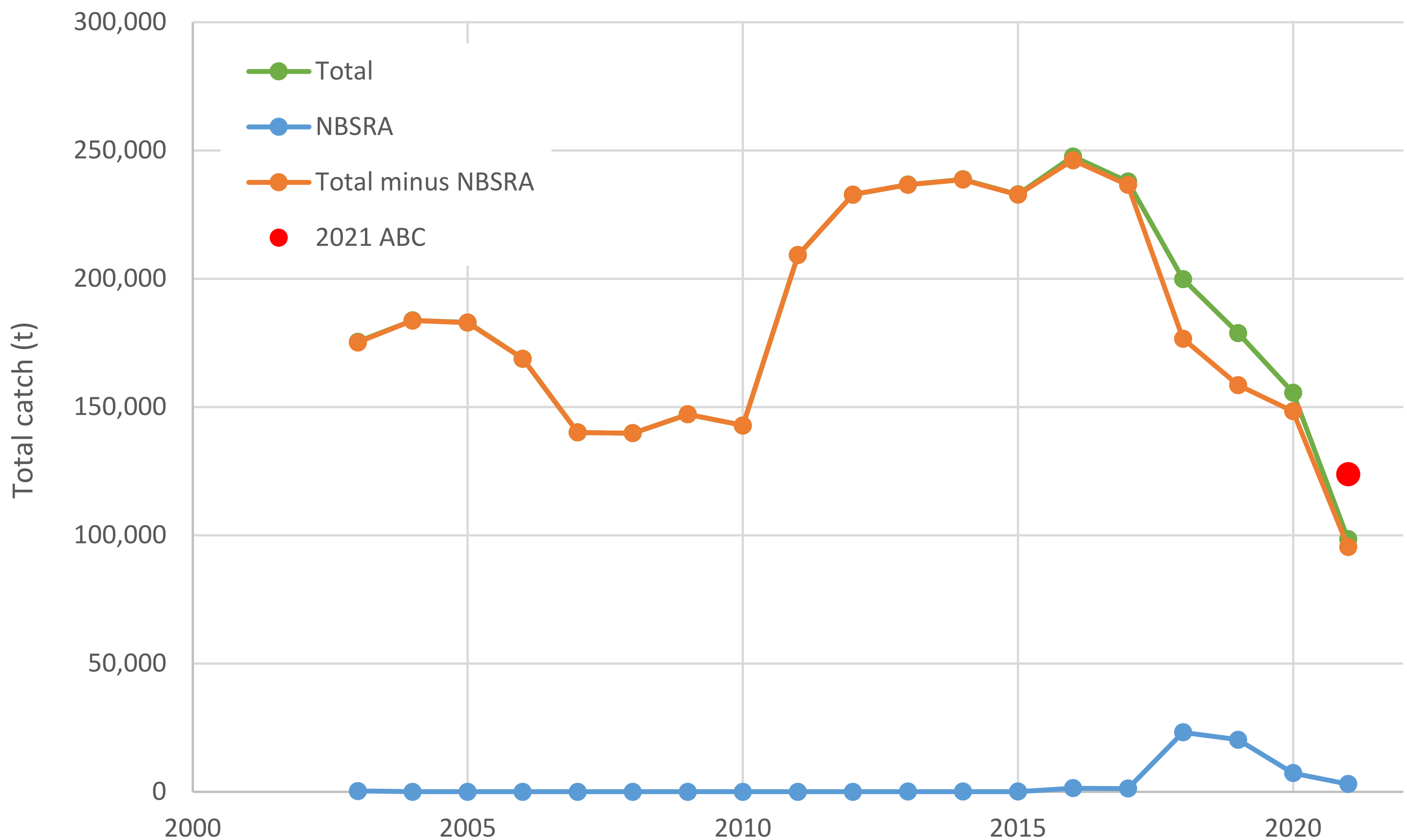
- Catch by gear (partial data for 2021)

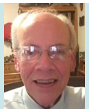




Data for context only: catch (2 of 2)

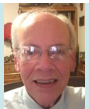
- Catch by area (partial data for 2021)





Data for context only: fishery CPUE (1 of 5)

- “Year effect” for 2021 is positive for all gear types
- Relative to the respective 1996-2020 monthly averages, highlights for months where data are available and can be reported:
  - Longline CPUE was > average from Jan-Aug except for May
  - Bottom trawl CPUE was > average from Feb-Apr
    - In particular, CPUE for Feb. was very high (> 3x average), repeating the performance observed in 2020
  - Pot CPUE was > average for Jan.
  - Pelagic trawl CPUE (including incidental catches) was mixed from Jan-Mar, with Jan. being < average and Feb-Mar being > average
    - In particular, CPUE for Mar. was > 2x average

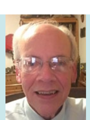


Data for context only: fishery CPUE (2 of 5)

- Longline, kg/hook (Pacific cod target only, normalized to mean=1.0)

Year	1	2	3	4	5	6	7	8	9	10	11	12
1996	1.646	1.665	1.480	1.366	1.263				0.986	1.040	1.086	
1997	1.807	1.908	1.597	1.576	1.242				1.081	1.085	1.033	1.142
1998	1.605	1.740	1.300	1.031	0.959				0.691	0.751	0.907	0.973
1999	1.361	1.431	1.202	1.090	1.232			0.817	0.960	0.863	0.974	1.086
2000	1.588	1.156	1.202	1.059	1.142			0.865	0.769	0.688	0.715	0.846
2001	1.118	1.073	1.065	0.949	0.943	1.052	0.865	0.782	0.729	0.734	0.724	0.878
2002	1.290	1.249	1.267	1.310			0.666	0.727	0.685	0.659	0.705	0.784
2003	0.933	0.984	1.071	0.864	0.838		0.651	0.643	0.632	0.629	0.651	0.750
2004	0.999	1.186	1.136	1.098	0.780		0.657	0.605	0.576	0.565	0.715	0.966
2005	1.168	1.193	1.272	1.261			0.711	0.640	0.580	0.620	0.641	0.815
2006	1.308	1.530	1.521	1.453			0.686	0.811	0.747	0.630	0.790	0.835
2007	1.356	1.406	1.339				0.711	0.873	0.729	0.649	0.807	1.185
2008	1.455	1.556	1.463	1.525			0.646	0.682	0.578	0.488	0.619	1.262
2009	1.632	1.795	2.194				0.650	0.713	0.646	0.625	0.685	1.034
2010	1.395	1.616	1.734				0.752	0.728	0.652	0.617	0.779	0.934
2011	1.287	1.393	1.390	1.248	0.851	0.821	0.611	0.652	0.683	0.725	0.767	0.909
2012	1.413	1.534	1.119	1.137	0.935	0.950	0.693	0.623	0.585	0.620	0.671	1.044
2013	1.424	1.377	1.257	1.234	0.994	0.688	0.766	0.689	0.652	0.647	0.800	1.023
2014	1.012	1.210	1.020	1.018	0.797	0.684	0.575	0.649	0.663	0.719	0.790	0.853
2015	0.983	1.197	1.125	1.017	0.952	0.804	0.847	0.772	0.655	0.699	0.788	0.961
2016	1.172	1.353	1.096	1.008	0.971	0.773	0.795	0.775	0.774	0.728	0.794	0.947
2017	1.022	1.399	1.220	1.130	0.977	0.856	0.713	0.574	0.595	0.668	0.898	1.044
2018	1.532	1.639	1.351	1.342	0.866	0.708	0.570	0.571	0.798	0.794	0.711	0.986
2019	1.626	1.792	1.383	1.369	0.907	0.905	0.678	0.679	0.791	0.844	0.930	0.974
2020	1.481	1.784	1.598	1.293	1.467	1.020	0.802	0.741	0.785	0.924	0.922	0.986
2021	1.364	1.569	1.532	1.247	0.735	1.043	0.941	0.823				

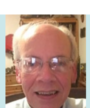
NOAA FISHERIES



Data for context only: fishery CPUE (3 of 5)

- Bottom trawl, kg/haul (Pacific cod target only, normalized to mean=1.0)

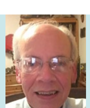
Year	1	2	3	4	5	6	7	8	9	10	11	12
1996	0.520	0.789	0.967	1.024		0.357					0.724	
1997	1.123	1.644	1.219	0.834	0.409	0.168	0.480	0.490				
1998	1.045	1.516	0.818	0.488	0.355			0.401	0.365	0.158	0.173	
1999	0.563	1.121	0.775	0.821	0.471				0.365	0.324		
2000	0.790	0.781	0.843	0.550	0.617	0.253	0.247					
2001	0.271	0.592	0.702	0.361	0.332	1.012	1.354	0.333	0.342	0.161		
2002	0.755	0.910	0.806	0.403	0.485	0.943	0.689	0.281	0.317	0.304	0.630	
2003	0.418	0.555	0.766	0.459	0.288	0.893	0.958	0.412	0.435			
2004	1.098	1.286	1.290	0.640	0.548	0.532	0.637	0.384	0.403	0.217		
2005	0.700	0.866	1.609	0.566	0.418	0.335	0.208	0.146				
2006	0.668	0.831	0.956	1.061	0.573		0.179	0.344	0.173			
2007	0.458	0.632	1.033	0.809	0.674	0.821	0.449		0.008			
2008	0.392	0.553	0.644	0.714	0.934	0.176		1.817	0.280			
2009	0.438	0.871	1.323	1.811	0.788	0.600		0.507		1.079		
2010	0.544	1.004	1.236		0.400		1.034	0.629		0.949		
2011	1.089	1.213	2.047	0.867	0.892		1.123	0.924	1.243	1.056		
2012	2.120	2.500	1.703	1.033	1.333	0.964		0.798	0.677			
2013	1.447	1.710	1.106	2.027	0.568	0.493		1.244	1.533			
2014	1.306	1.302	1.513	1.296	1.519			0.763	1.270			
2015	0.438	1.376	1.777	1.180	1.464		1.288	2.590				
2016	0.923	1.543	2.133	1.140		0.588	0.961	0.618	0.558	0.735		
2017	0.982	2.124	2.146	1.899		0.386			0.132			
2018	1.394	1.646	6.402	5.496								
2019	1.351	1.763	3.977	6.247	0.187				0.364			
2020	0.687	4.678	1.575	0.710	0.440							
2021		4.616	2.813	1.962								



Data for context only: fishery CPUE (4 of 5)

- Pot, kg/pot (Pacific cod target only, normalized to mean=1.0)

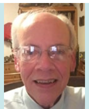
Year	1	2	3	4	5	6	7	8	9	10	11	12
1996			0.929				0.593	0.614	0.516	0.576		
1997			1.066	1.245	0.682	0.636	0.524	0.689	0.379	0.700	0.493	0.545
1998				1.005	0.682	0.584	0.689	0.473	0.406	0.370		0.658
1999				0.880	0.621	0.486			0.427	0.512		
2000		0.743	0.593									
2001			0.839						0.574	0.675	0.419	
2002		1.110	0.853						0.559	0.538	0.635	
2003		0.970	0.904						0.742	0.809	0.777	
2004	0.805	1.035	0.936						0.632	0.522	0.873	
2005	1.439	1.062							0.675	0.521		
2006	1.103	0.945							0.632	0.869	0.454	0.492
2007	0.816								0.815	1.203		
2008	1.056								0.828	0.548		
2009	1.189											
2010	1.114								1.111	0.953	1.358	
2011	1.704								1.086	0.988		
2012	1.866								0.742	1.055		
2013	1.635	1.728							1.253	1.043	1.150	
2014	1.244	1.565	2.079						1.132	0.942	0.707	1.334
2015	1.193	1.751	1.796	1.699					1.207	1.035	0.853	1.122
2016	1.152	1.421	1.505	1.443					1.302	1.035	0.843	1.054
2017	1.205	2.048							0.918	0.929	0.911	0.839
2018	1.162		2.356						1.156	1.324	0.720	1.146
2019	1.232		2.788						1.764		1.718	
2020	1.224								1.742			
2021	1.446											



Data for context only: fishery CPUE (5 of 5)

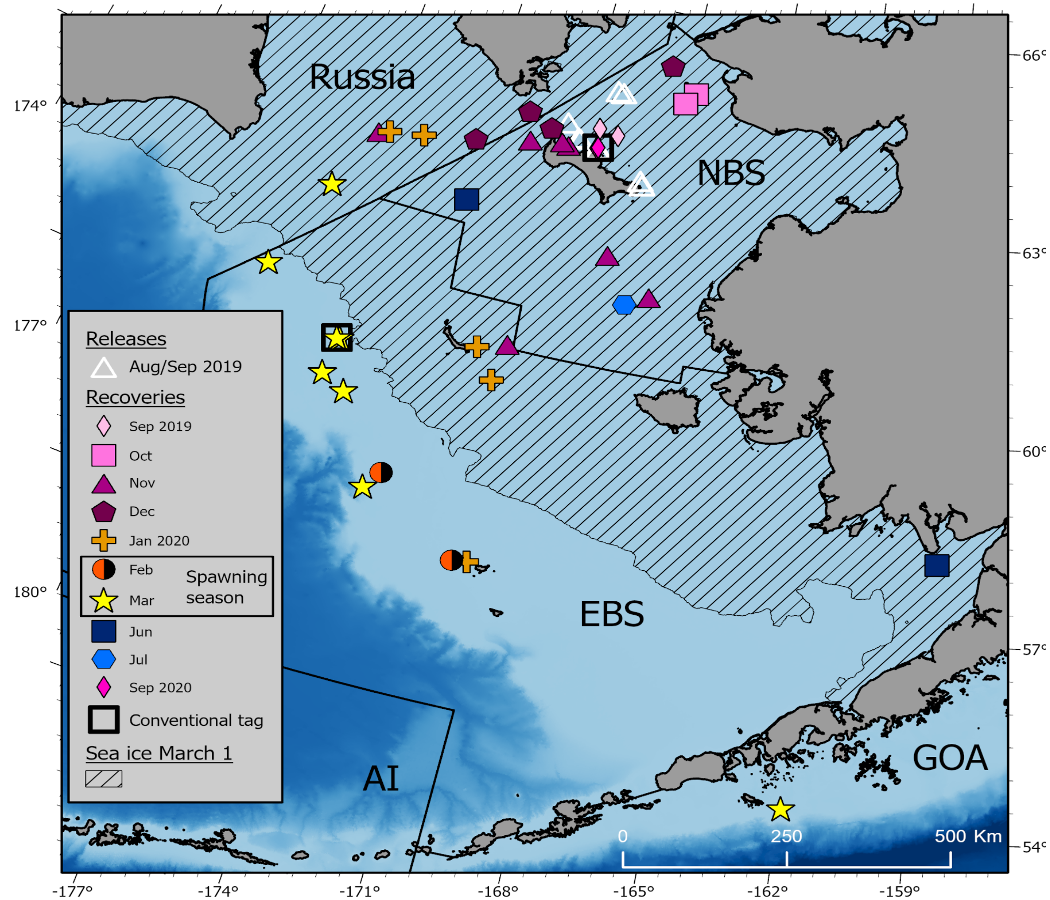
- Pelagic trawl, kg/haul (*all* hauls, normalized to mean=1.0)

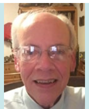
Year	1	2	3	4	5	6	7	8	9	10	11	12
1996	2.432	1.167	0.842					0.263	0.530	0.436	0.603	
1997	4.976	1.911	2.422	2.522				0.381	0.514	0.464		
1998	2.558	1.384	4.247					0.130	0.430	0.677	0.459	
1999	1.570	0.964	0.623				0.389	0.441	0.412	0.395	0.228	
2000	4.365	0.736	0.654	0.336			0.283	0.248	0.379	0.298	0.629	
2001	1.272	0.595	0.456	0.621		0.225	0.294	0.481	0.286	0.335	0.116	
2002	2.036	1.682	0.982	1.691		0.280	0.250	0.366	0.465	0.416		
2003	3.236	1.493	0.772	1.301		0.272	0.278	0.357	0.432	0.306		
2004	1.978	1.884	0.968			0.495	0.254	0.231	0.376	0.202		
2005	2.619	1.522	1.541			0.296	0.244	0.279	0.449	0.309		
2006	2.649	1.673	1.428			0.304	0.354	0.459	0.349	0.295	0.443	
2007	1.003	1.054	1.272			0.307	0.407	0.405	0.269	0.288	0.339	
2008	1.204	1.002	1.674			0.604	0.424	0.378	0.206	0.141	0.078	
2009	1.080	1.467	1.648	3.581		0.298	0.451	0.383	0.211	0.290		
2010	1.241	1.828	1.523	2.103		0.481	0.531	0.326	0.361	0.323		
2011	1.379	1.735	1.472	1.164		0.417	0.378	0.249	0.253	0.332	0.411	
2012	3.047	2.766	1.681	0.830		0.327	0.617	0.264	0.331	0.409	0.411	
2013	1.099	1.275	1.765	1.479		0.682	0.512	0.454	0.349	0.427		
2014	0.731	0.600	0.739	2.210		0.354	0.345	0.371	0.421			
2015	0.613	1.519	1.729	1.962		0.485	0.839	0.812	0.964	1.260		
2016	0.651	1.142	1.218	2.102		0.644	0.437	0.330	0.253	0.225		
2017	1.050	1.487	1.973	4.658		0.659	0.387	0.367	0.311	1.940		
2018	0.596	2.026	1.758	2.362		0.360	0.133	0.162	0.199			
2019	1.810	2.870	3.096	4.441		0.390	0.204	0.196	0.127	0.064		
2020	0.995	3.646	3.290	2.955	2.764	0.200	0.196	0.292	0.216	0.141		
2021	1.333	1.827	3.571	2.240	2.196	0.664						



Data for context only: tagging (1 of 2)

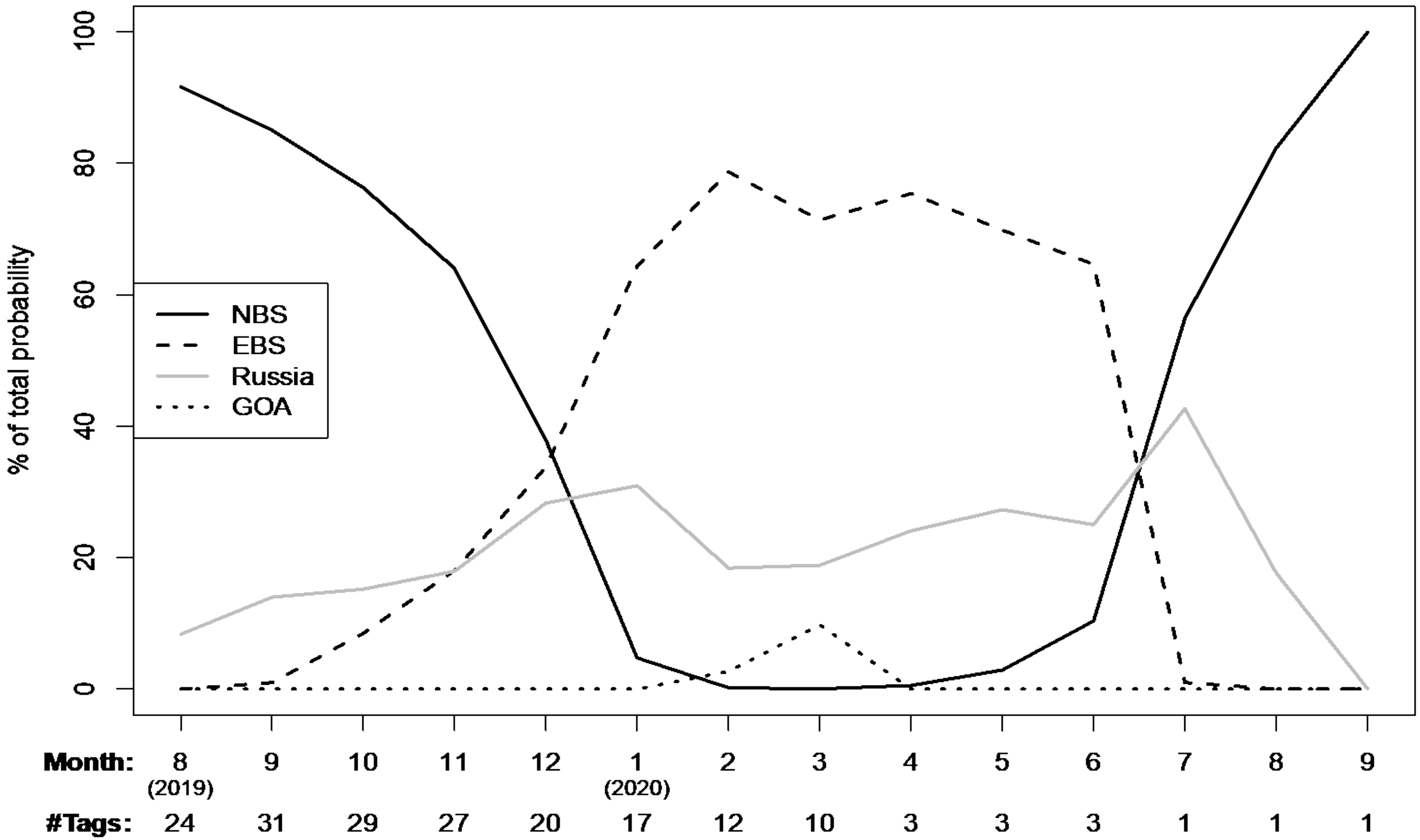
- Release (NBS) and recovery locations





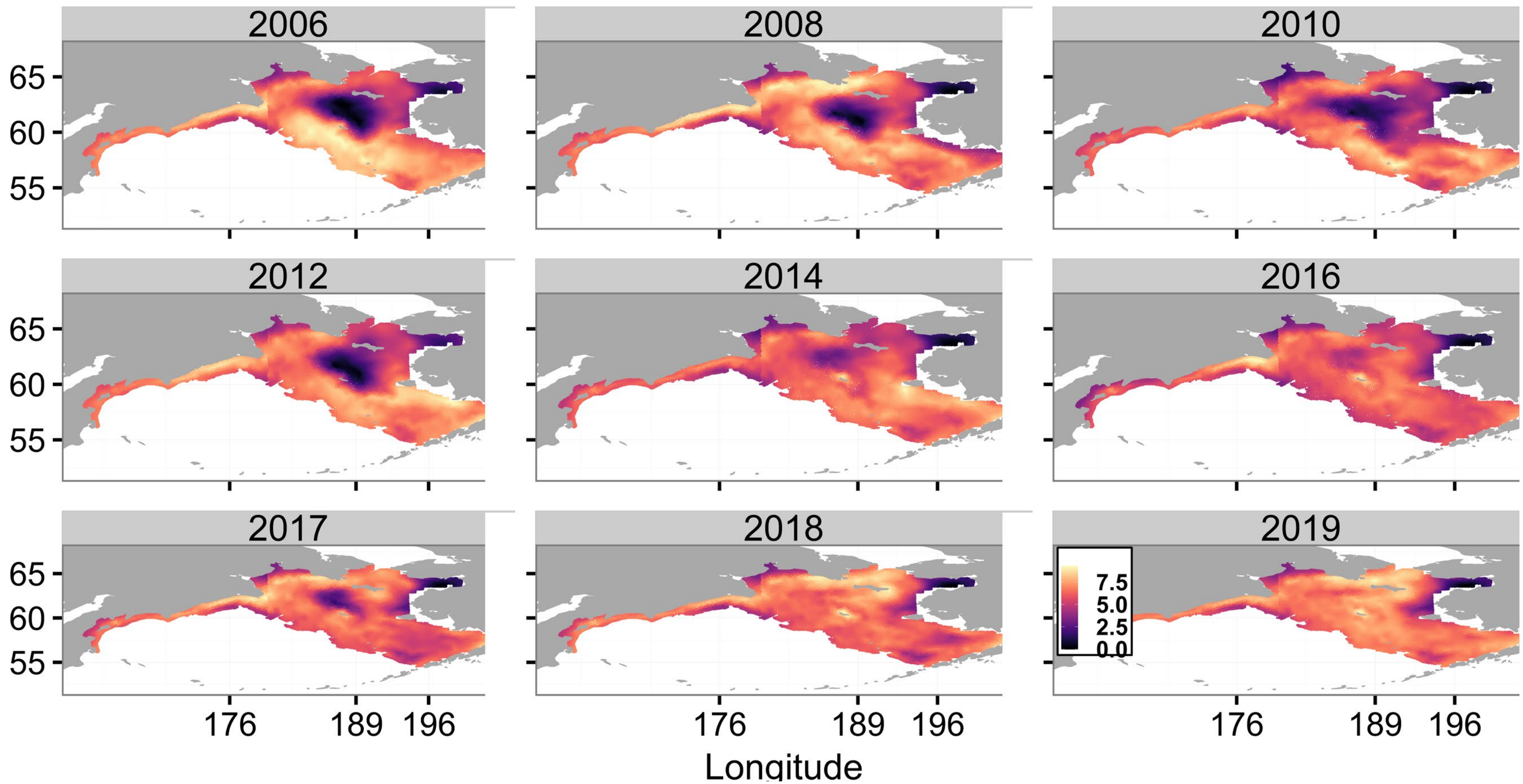
Data for context only: tagging (2 of 2)

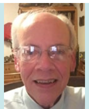
- Monthly location probability



Data for context only: Bering-wide surveys

- VAST density estimates based on WBS, EBS, and NBS surveys





“New” data used: VAST survey index (1 of 8)

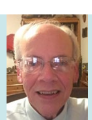
- In response to comment SSC1, alternative configurations for the use of covariates, number of knots, and error distribution were explored
- VAST estimates were obtained for each of the following configurations:

No.	Cold pool covariate?	Knots	Distribution
1	yes	750	Poisson-linked delta-gamma
2	no	750	Poisson-linked delta-gamma
3	yes	100	Poisson-linked delta-gamma
4	yes	750	Tweedie

- Configuration #1 was chosen on the basis of AIC:

No.	Cold pool?	Knots	Distribution	AIC	ΔAIC
1	Yes	750	P-link Δ-gamma	220372.7	0
2	No	750	P-link Δ-gamma	220480.6	107.9
3	Yes	100	P-link Δ-gamma	222030.2	1657.5
4	Yes	750	Tweedie	222268.2	1895.5

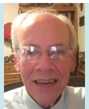
NOAA FISHERIES



“New” data used: VAST survey index (2 of 8)

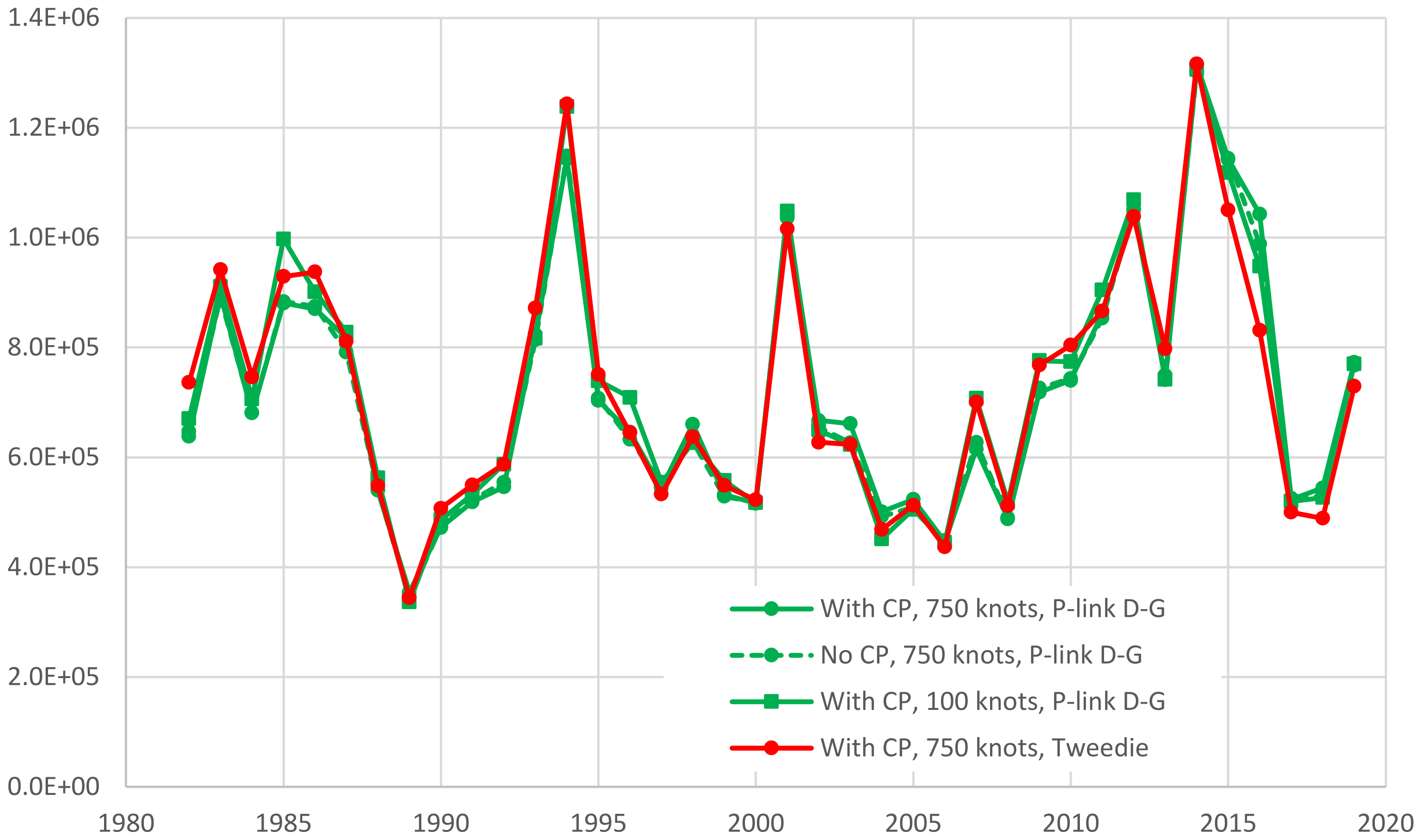
- Correlations

Area	Cold pool?	Knots	Distribution				
	Cold pool?			Yes	No	Yes	Yes
		Knots		750	750	100	750
Area			Distribution	P-link Δ-gamma	P-link Δ-gamma	P-link Δ-gamma	Tweedie
EBS+NBS	Yes	750	P-link Δ-gamma	1.000	0.998	0.985	0.968
EBS+NBS	No	750	P-link Δ-gamma	0.998	1.000	0.990	0.977
EBS+NBS	Yes	100	P-link Δ-gamma	0.985	0.990	1.000	0.987
EBS+NBS	Yes	750	Tweedie	0.968	0.977	0.987	1.000
EBS	Yes	750	P-link Δ-gamma	1.000	1.000	0.991	0.988
EBS	No	750	P-link Δ-gamma	1.000	1.000	0.992	0.989
EBS	Yes	100	P-link Δ-gamma	0.991	0.992	1.000	0.993
EBS	Yes	750	Tweedie	0.988	0.989	0.993	1.000
NBS	Yes	750	P-link Δ-gamma	1.000	0.992	0.982	0.898
NBS	No	750	P-link Δ-gamma	0.992	1.000	0.983	0.901
NBS	Yes	100	P-link Δ-gamma	0.982	0.983	1.000	0.946
NBS	Yes	750	Tweedie	0.898	0.901	0.946	1.000

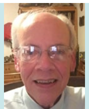


“New” data used: VAST survey index (3 of 8)

- Survey abundance (1000s of fish, EBS and NBS combined)

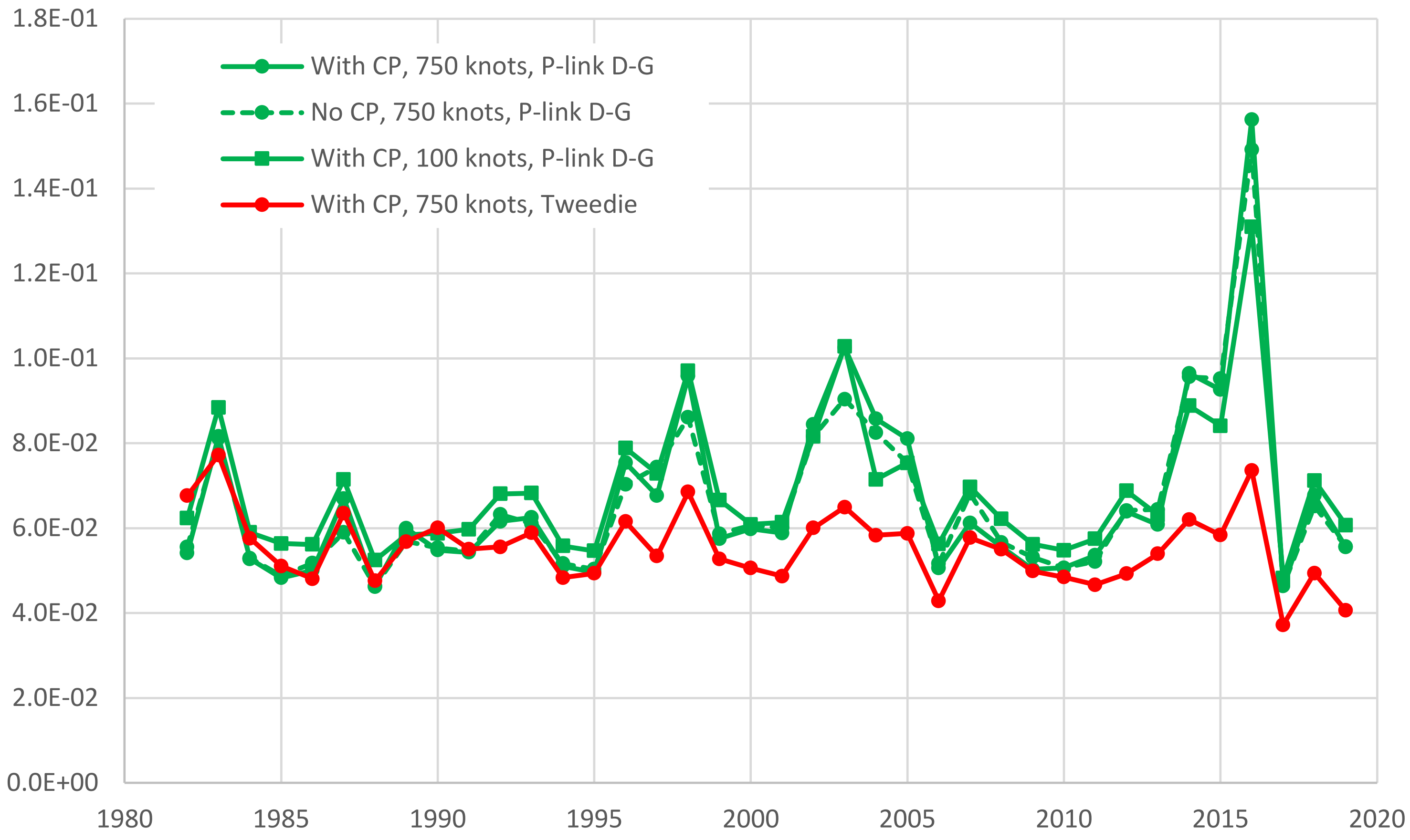


NOAA FISHERIES



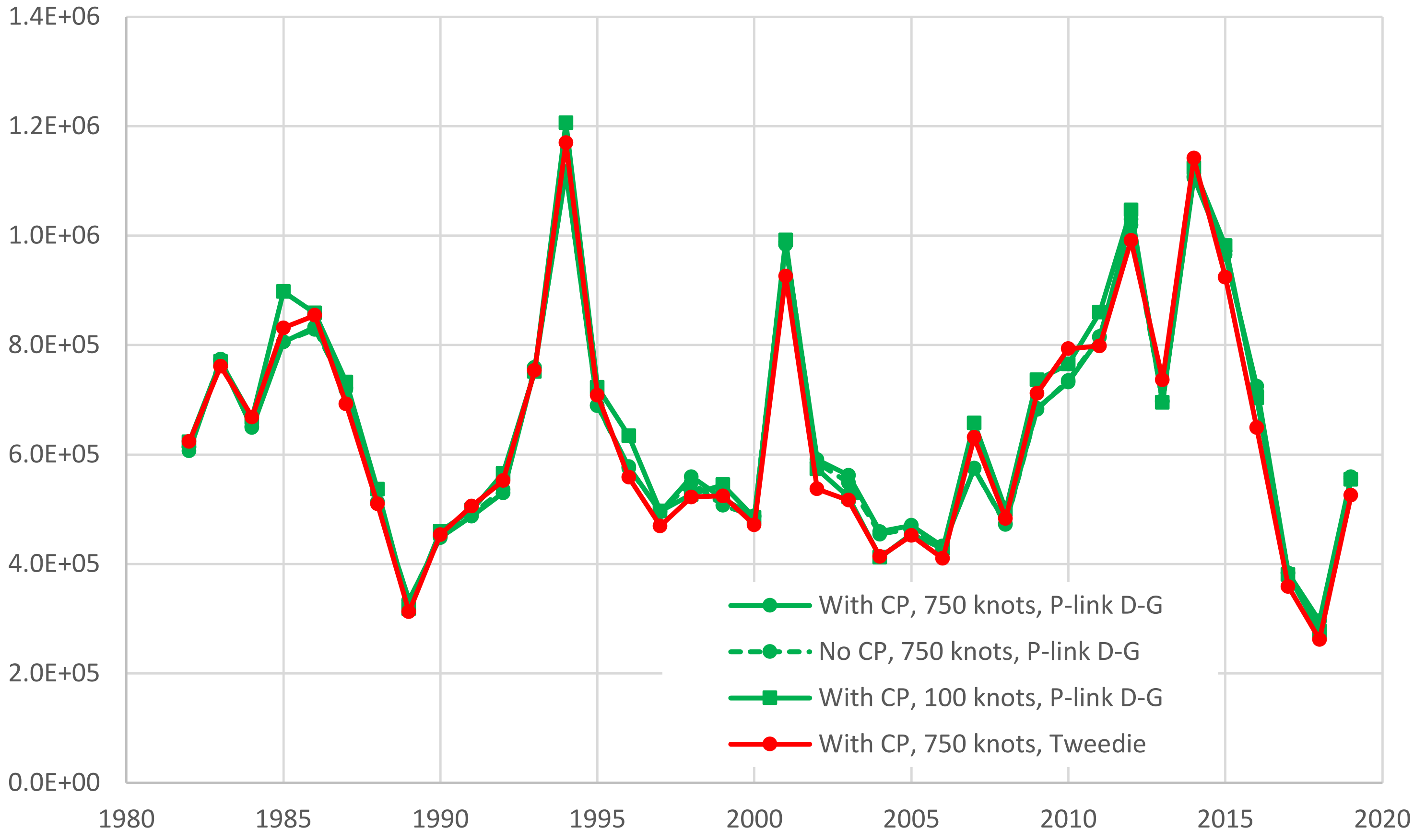
“New” data used: VAST survey index (4 of 8)

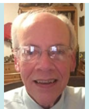
- Lognormal sigma (EBS and NBS combined)



“New” data used: VAST survey index (5 of 8)

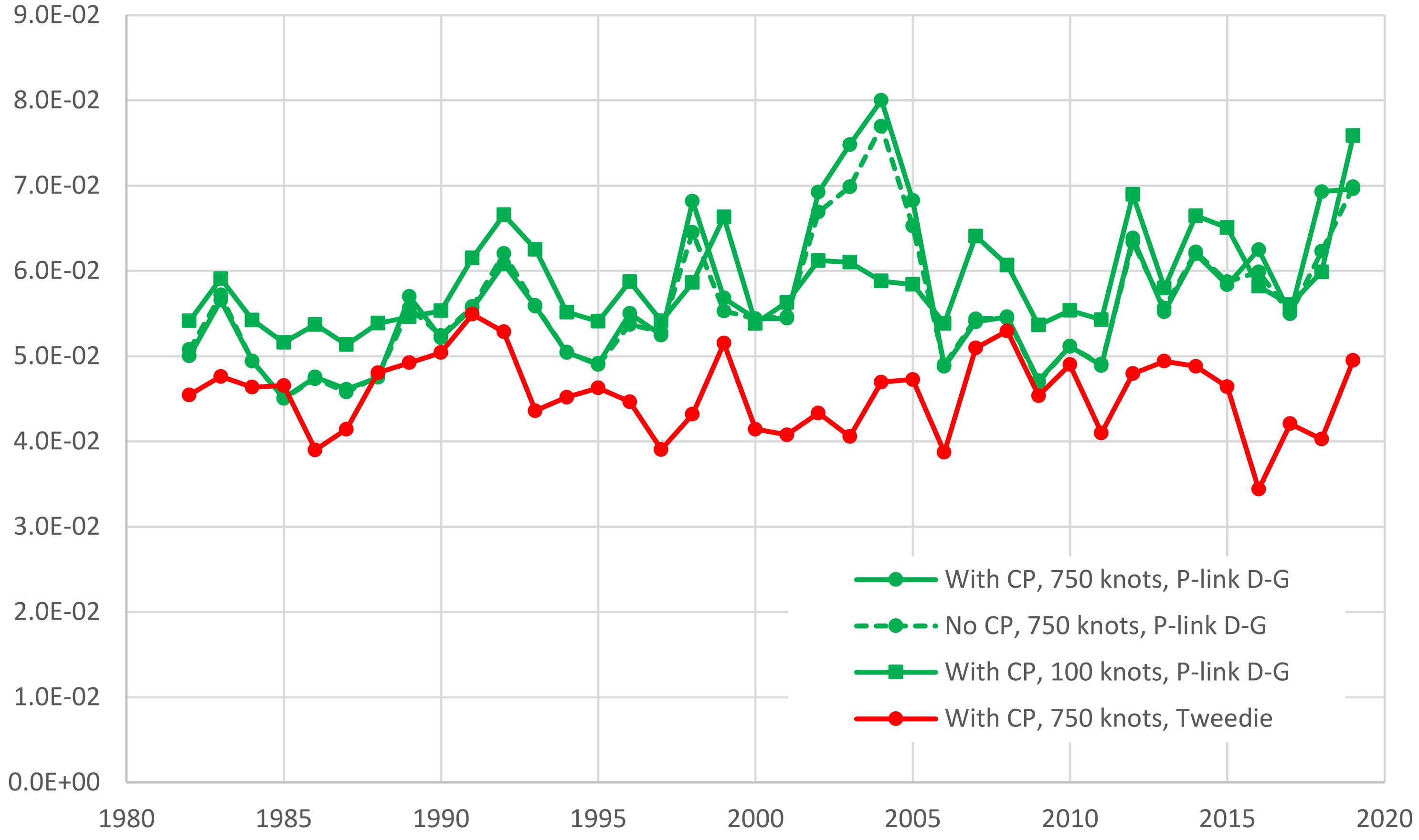
- Survey abundance (EBS only)

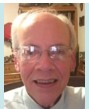




“New” data used: VAST survey index (6 of 8)

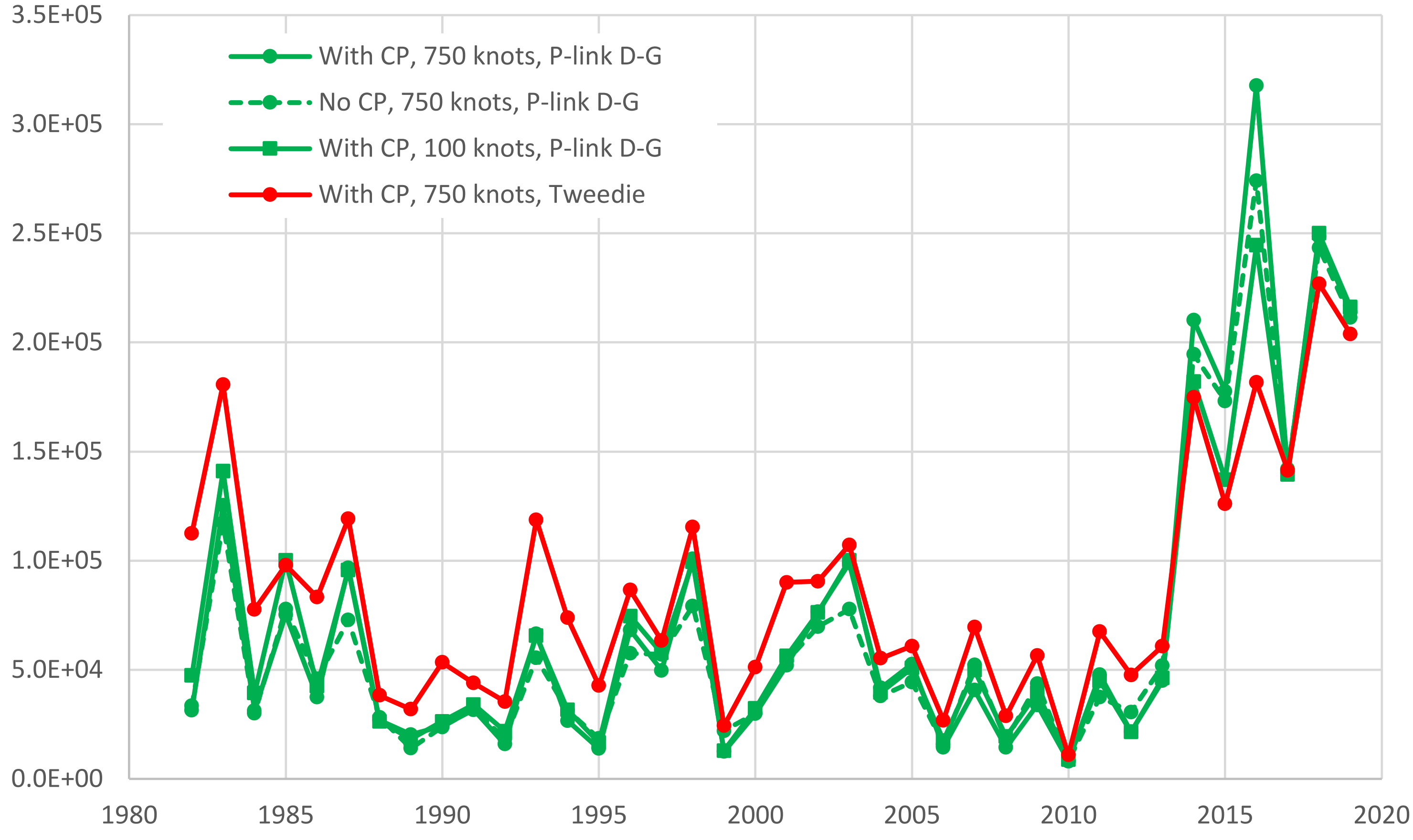
- Lognormal sigma (EBS only)

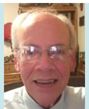




“New” data used: VAST survey index (7 of 8)

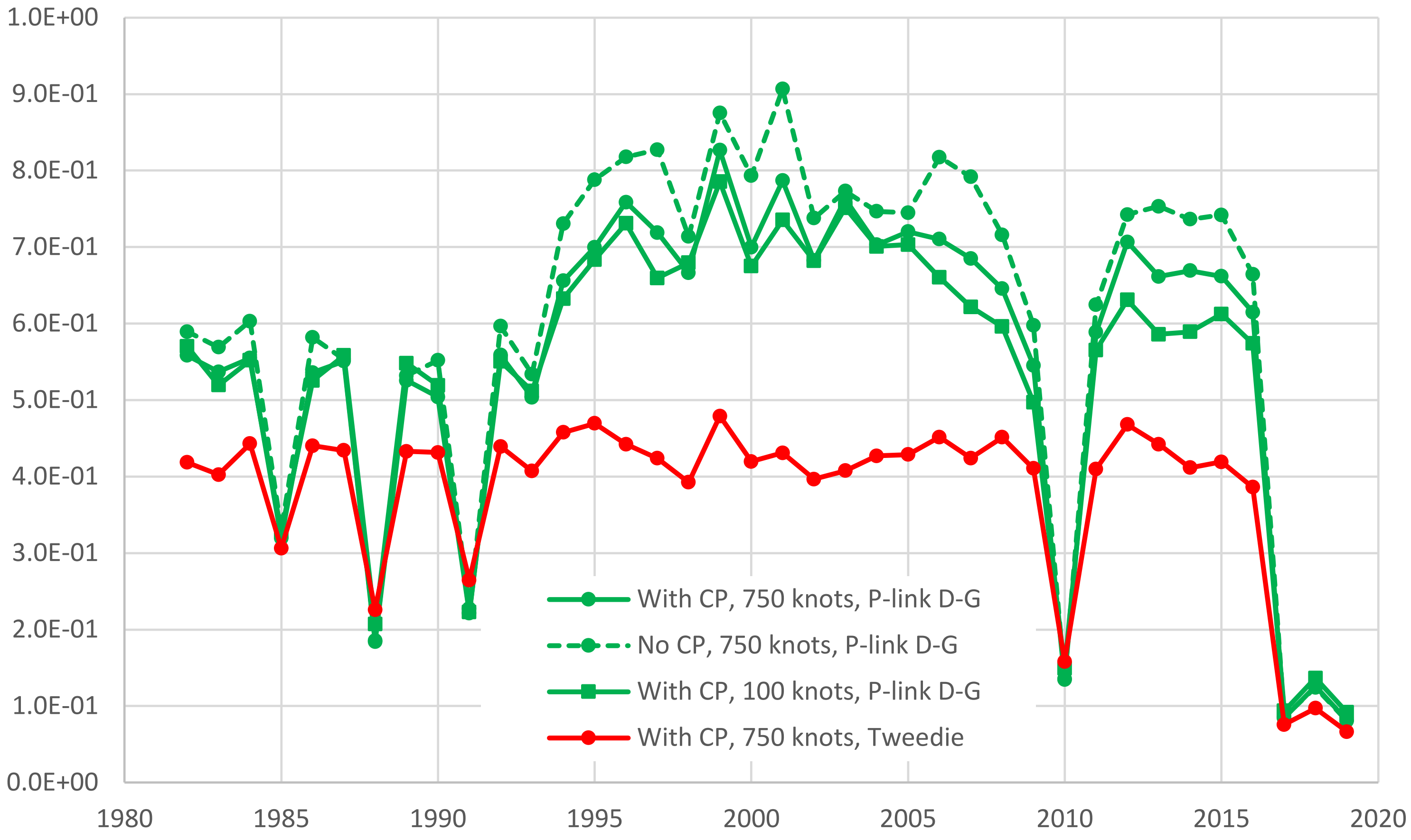
- Survey abundance (NBS only)





“New” data used: VAST survey index (8 of 8)

- Lognormal sigma (NBS only)

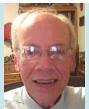


“New” data used: VAST agecomps

- Edits to the underlying data set necessitated re-running VAST

Year	0	1	2	3	4	5	6	7	8	9	10	11	12+
1994	0.00025	0.10960	0.36178	0.16724	0.11357	0.11841	0.08868	0.02324	0.00894	0.00437	0.00148	0.00115	0.00128
1995	0.00016	0.06394	0.24421	0.41964	0.10467	0.07786	0.05798	0.01463	0.00731	0.00548	0.00141	0.00142	0.00128
1996	0.00003	0.06777	0.18147	0.17359	0.28545	0.15581	0.08064	0.03683	0.00964	0.00364	0.00189	0.00163	0.00160
1997	0.00029	0.27106	0.16713	0.14830	0.14075	0.12328	0.10537	0.02746	0.01106	0.00218	0.00172	0.00081	0.00058
1998	0.00007	0.07735	0.42642	0.19449	0.11092	0.06057	0.06975	0.03585	0.01884	0.00370	0.00078	0.00082	0.00043
1999	0.00010	0.10650	0.18680	0.29143	0.21182	0.07817	0.06794	0.03386	0.01407	0.00600	0.00119	0.00144	0.00067
2000	0.00000	0.20120	0.10995	0.15719	0.23899	0.17274	0.07443	0.01709	0.01716	0.00498	0.00393	0.00167	0.00068
2001	0.00004	0.27894	0.22734	0.17678	0.08903	0.09894	0.08474	0.03119	0.00801	0.00185	0.00160	0.00110	0.00045
2002	0.00023	0.07471	0.17881	0.29519	0.24556	0.07753	0.06866	0.04348	0.01062	0.00309	0.00103	0.00049	0.00060
2003	0.00001	0.16659	0.14230	0.23106	0.21247	0.13663	0.05220	0.03584	0.01724	0.00374	0.00051	0.00063	0.00079
2004	0.00005	0.13003	0.15148	0.26514	0.12787	0.13909	0.10903	0.04254	0.02159	0.00817	0.00225	0.00196	0.00080
2005	0.00000	0.14653	0.22804	0.20325	0.12604	0.07255	0.10355	0.07164	0.02821	0.01106	0.00411	0.00445	0.00057
2006	0.00000	0.33628	0.13637	0.15493	0.10952	0.09310	0.06940	0.05420	0.02987	0.01053	0.00353	0.00137	0.00090
2007	0.00000	0.66394	0.09784	0.07039	0.04857	0.05437	0.02432	0.02031	0.00992	0.00602	0.00207	0.00120	0.00105
2008	0.00000	0.21988	0.41153	0.14618	0.08927	0.05506	0.03663	0.01380	0.01322	0.00739	0.00320	0.00231	0.00154
2009	0.00000	0.48663	0.17650	0.21229	0.05935	0.02726	0.01581	0.01123	0.00611	0.00222	0.00127	0.00082	0.00052
2010	0.00000	0.05070	0.49801	0.17026	0.18543	0.06049	0.01713	0.01047	0.00416	0.00185	0.00070	0.00064	0.00016
2011	0.00008	0.30585	0.06959	0.36466	0.11080	0.09755	0.03340	0.00929	0.00408	0.00197	0.00139	0.00077	0.00056
2012	0.00000	0.36834	0.24105	0.05898	0.21603	0.06419	0.03542	0.00961	0.00306	0.00213	0.00073	0.00018	0.00027
2013	0.00000	0.10458	0.35269	0.19492	0.12110	0.13941	0.06485	0.01571	0.00449	0.00124	0.00029	0.00037	0.00035
2014	0.00004	0.28420	0.17128	0.22332	0.19972	0.05718	0.04679	0.01231	0.00252	0.00099	0.00093	0.00010	0.00061
2015	0.00002	0.05978	0.41467	0.20764	0.19774	0.08474	0.01977	0.01199	0.00257	0.00049	0.00027	0.00011	0.00020
2016	0.00000	0.08644	0.09400	0.35140	0.22860	0.16502	0.05655	0.01214	0.00359	0.00136	0.00049	0.00027	0.00014
2017	0.00007	0.10561	0.17242	0.15945	0.29978	0.15092	0.08353	0.02037	0.00305	0.00299	0.00061	0.00053	0.00067
2018	0.00003	0.07259	0.09656	0.25352	0.16826	0.28695	0.08715	0.02951	0.00258	0.00174	0.00051	0.00022	0.00038
2019	0.00001	0.58946	0.07873	0.08517	0.07615	0.05909	0.07176	0.03172	0.00630	0.00082	0.00032	0.00023	0.00025

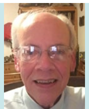
- Changes in proportions at age were generally small (never > 4%)



“New” data used: VAST fishery CPUE (1 of 5)

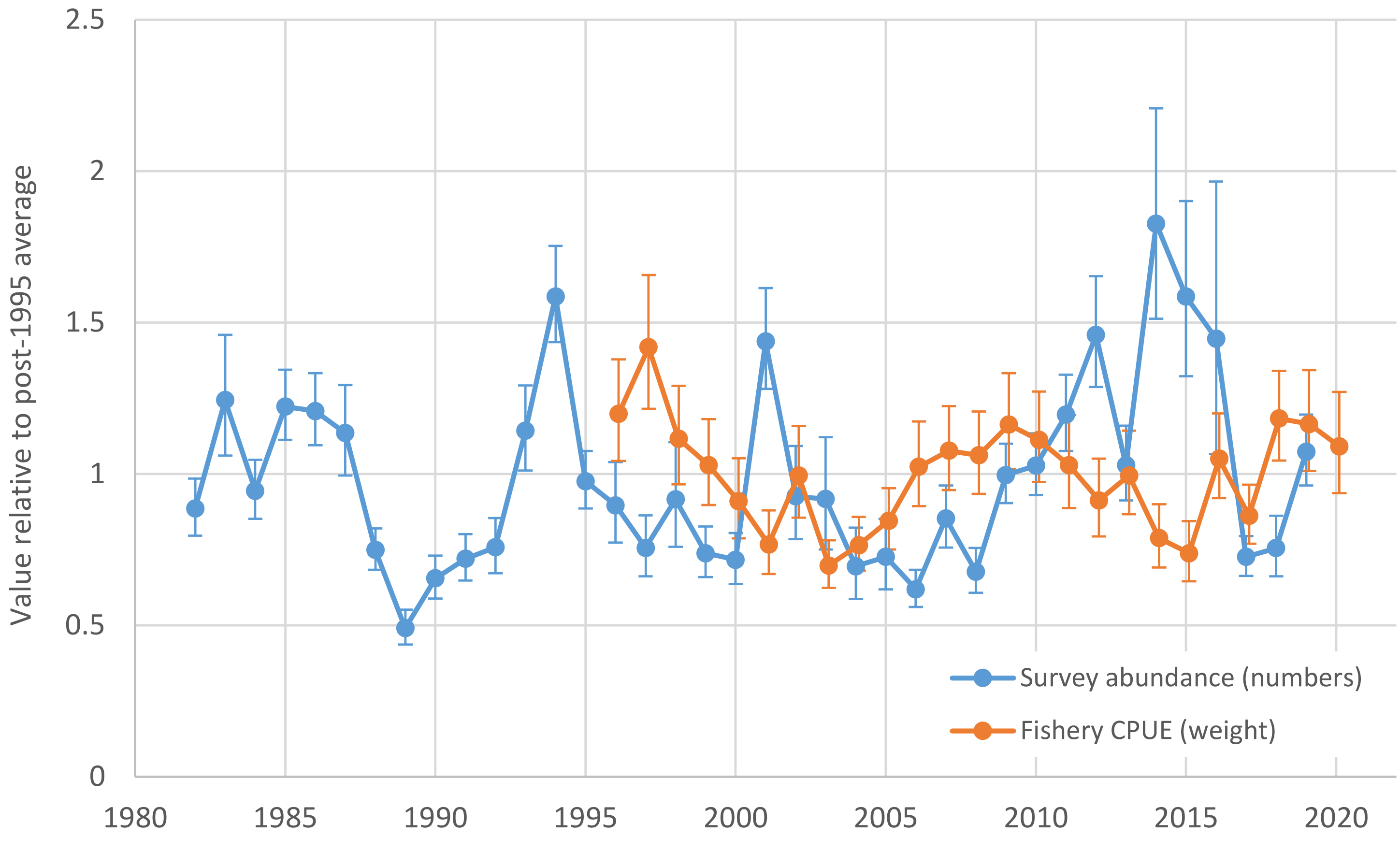
- Fishery CPUE index used in Model 20.9 was “of concern” to the Team
  - Team said to pursue “alternative statistical methods”
  - SSC said to develop “spatiotemporal analysis of fleet-specific CPUE”
- VAST CPUE index developed from catch (in weight) and effort data from January-February longline fishery
  - Jan-Feb chosen because fishery is within the EBS survey footprint
- Initially envisioned as just the first step in developing a CPUE index
  - That is, as a starting point for the CIE reviewers’ deliberations
- CIE reviewers recommended using it in one of this year’s models

NOAA FISHERIES



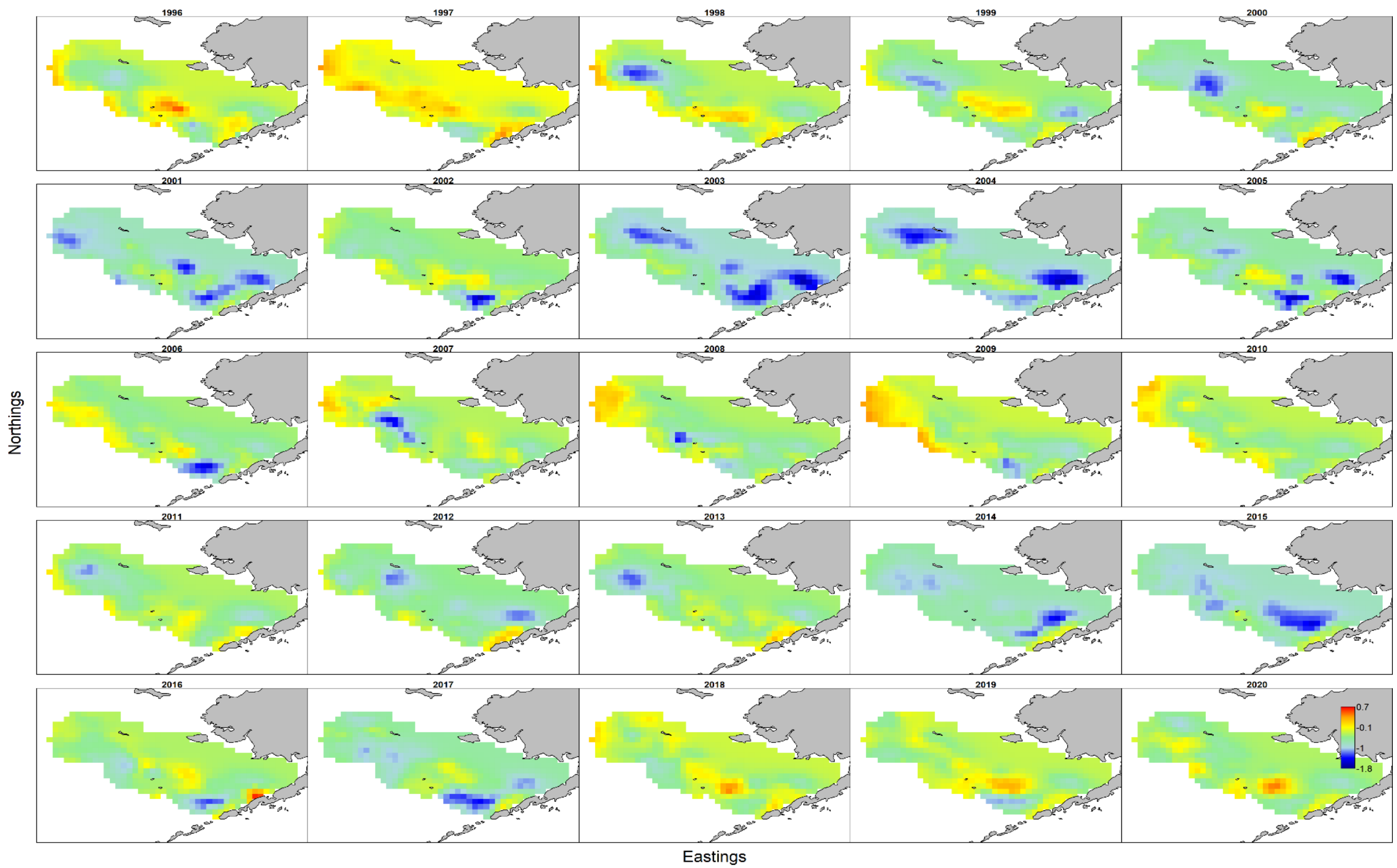
“New” data used: VAST fishery CPUE (2 of 5)

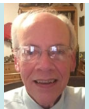
- Comparison of the fishery and survey indices (correlation = -0.338)



“New” data used: VAST fishery CPUE (3 of 5)

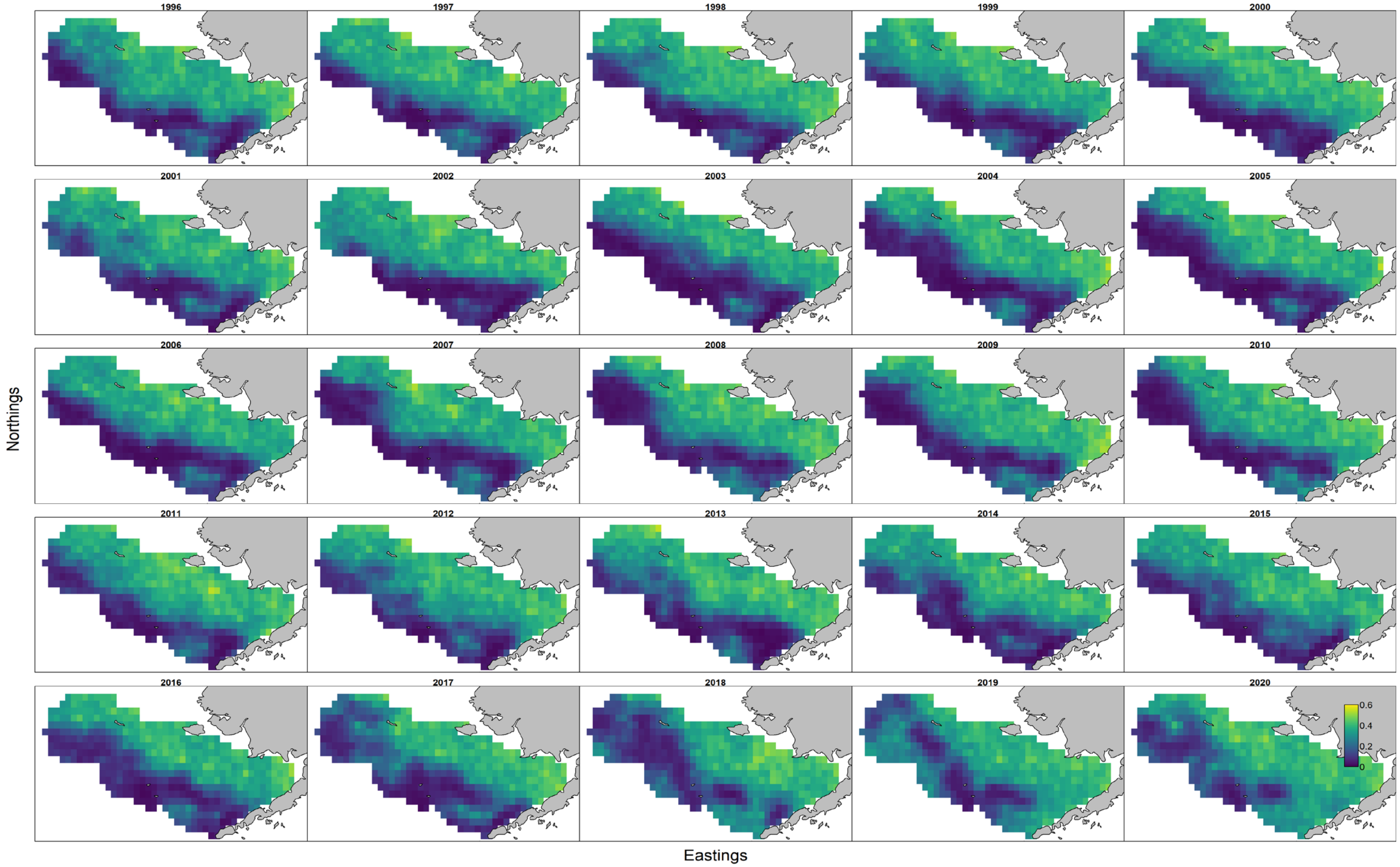
- Log density by year (1996-2020); blue=low, red=high

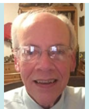




“New” data used: VAST fishery CPUE (4 of 5)

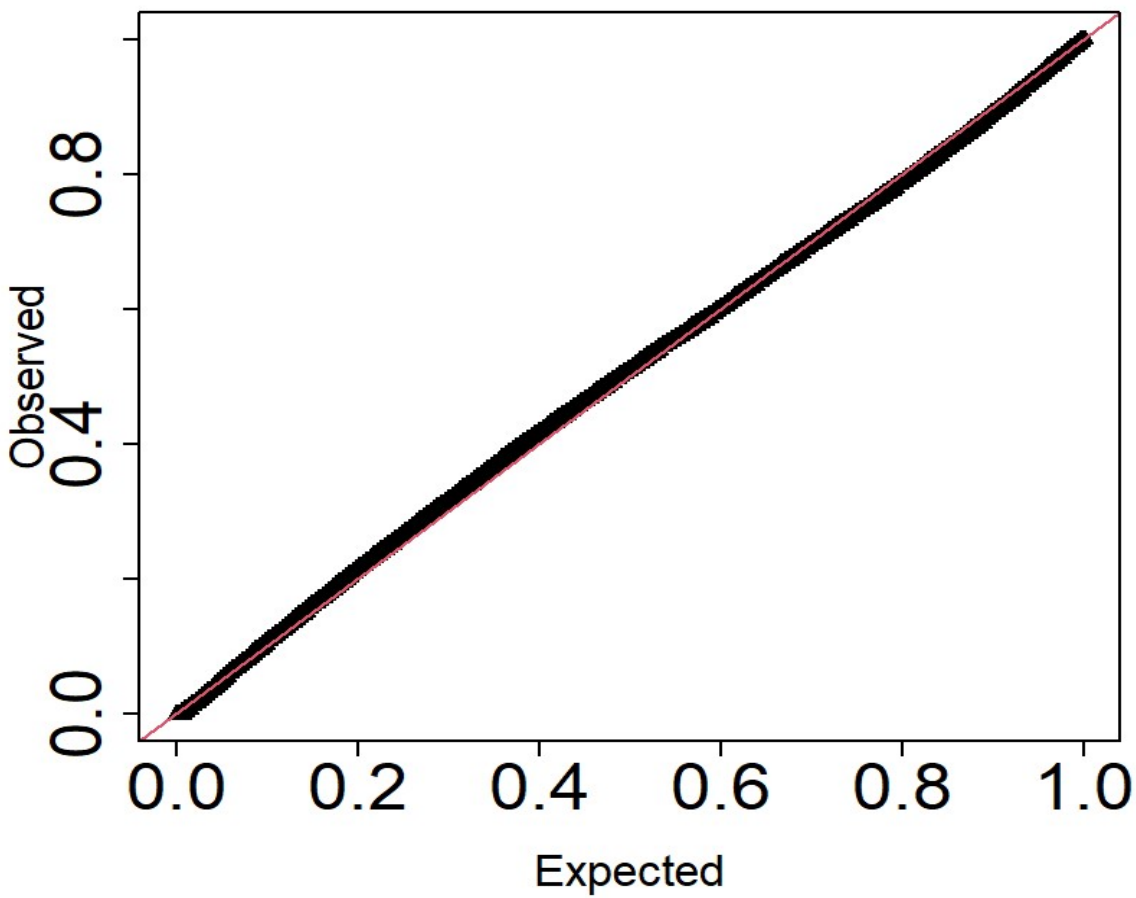
- Log density standard errors by year (1996-2020)



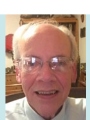


“New” data used: VAST fishery CPUE (5 of 5)

- Quantile-quantile plot



NOAA FISHERIES

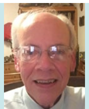


Models

The CIE model set (reprise)

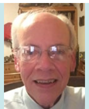
- Reviewers adopted a “+1” approach to specify the set of models
  - Base model = Model 19.12a
  - Four new features
  - Each new model = base model + 1 new feature

Feature	19.12a	19.12	"20.8a"	"20.9a"	"21.cie"
Feature 1: Allow catchability to vary?	no	yes	no	no	no
Feature 2: Allow domed survey selectivity?	no	no	yes	no	no
Feature 3: Use fishery CPUE?	no	no	no	yes	no
Feature 4: Estimate survey CV internally?	no	no	no	no	yes



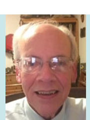
Base model (1 of 5)

- Sexes combined
- One season per year
- Natural mortality (constant across age and time) freely estimated
- Mean length at age follows a Richards growth function:
    - Base value of length at age 1.5 freely estimated
        - With constrained annual deviations on the log scale
    - Von Bertalanffy (Brody) growth coefficient freely estimated
    - Asymptotic length freely estimated
    - Richards growth coefficient freely estimated
- SD of L_at_A varies linearly with L_at_A, parameters freely estimated
- Weight at length varies annually, estimated outside the model
- Maturity at length (constant across time) estimated outside the model



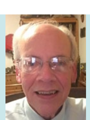
Base model (2 of 5)

- Mean ageing error varies with age, freely estimated within each block:
  - 1977-2007
  - 2008-present
- Recruitment is independent of stock size:
  - Mean freely estimated within each block:
    - Pre-1977
    - 1977-present
  - With constrained annual deviations on the log scale



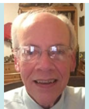
Base model (3 of 5)

- One survey, covering the EBS and NBS combined
  - Base value of log catchability freely estimated
  - Size-based, double-normal selectivity, with parameters as follow:
    - Base value of first size with selectivity=1 freely estimated
      - With constrained annual deviations on the log scale
    - Logit of size range with selectivity=1 fixed at 10.0
    - Base value of log of SD for 1st normal pdf freely estimated
      - With constrained annual deviations
    - Log of SD for 2nd normal pdf fixed at 10.0
    - Logit of selectivity at minimum size fixed at -10.0
    - Logit of selectivity at maximum size fixed at 10.0



Base model (4 of 5)

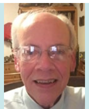
- One fishery, covering the EBS and NBS combined
  - Size-based, double-normal selectivity, with parameters as follow:
    - First size with selectivity=1 freely estimated
    - Logit of size range with selectivity=1 freely estimated
    - Base value of log of SD for 1st normal pdf freely estimated
      - With constrained annual deviations
    - Log of standard deviation for 2nd normal pdf freely estimated
    - Logit of selectivity at minimum size fixed at -10.0
    - Base value of logit of selectivity at maximum size freely estimated
      - With constrained annual deviations



Base model (5 of 5)

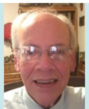
- Input sample sizes ($Nsamp$) for compositional data range between zero and an initial number ($Ninit$) according to the formula $Nsamp = (1 + \exp(ln\theta)\, Ninit)/(1 + \exp(ln\theta))$, where $ln\theta$ is a time-invariant parameter (the "Dirichlet-multinomial" parameter, estimated in natural log space, so that $Nsamp$ approaches 1 as $ln\theta$ approaches $-\infty$, $Nsamp = (1+Ninit)/2$ when $ln\theta = 0$, and $Nsamp$ approaches $Ninit$ as $ln\theta$ approaches $+\infty$), freely estimated for each of the compositional data types (fishery size composition data, survey size composition data, and survey age composition data), where:
  - For survey compositional data, $Ninit$ is the number of sampled hauls
  - For fishery compositional data, $Ninit$ is equal to the number of sampled hauls rescaled so that the average $Ninit$ for the fishery is equal to the average $Ninit$ for the survey (so that, on average, fishery data are emphasized equally with survey data)

NOAA FISHERIES



Tuning of deviation vector “σ” terms

- For the vector of deviations associated with log catchability (Model 19.12 only), σ was tuned to set RMSSR=1.0
- For the vector of deviations associated with log-scale recruitment, σ was tuned to match the square root of the variance of the estimates plus the sum of the estimates’ variances (Methot and Taylor 2011)
- For all other vectors of deviations, σ was tuned to set the variance of the estimates plus the sum of the estimates’ variances equal to 1.0



Structural comparison of models (1 of 2)

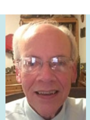
- Focusing on parameter deviations:

Model	19.12a	19.12	"20.8a"	"20.9a"	"21.cie"
Feature 1: Allow catchability to vary?	no	yes	no	no	no
Feature 2: Allow domed survey selectivty?	no	no	yes	no	no
Feature 3: Use fishery CPUE?	no	no	no	yes	no
Feature 4: Estimate survey CV internally?	no	no	no	no	yes
"Early" recruitment deviations	20	20	20	20	20
"Main" recruitment deviations	43	43	43	43	43
Length at age 1.5 deviations	43	43	43	43	43
Selectivity (fishery) deviations	88	88	88	88	88
Selectivity (survey) deviations	76	76	76	76	76
Log catchability (survey) deviations		38			
Annual deviations	270	308	270	270	270

Structural comparison of models (2 of 2)

- Focusing on true (unconstrained) parameters:

Model	19.12a	19.12	"20.8a"	"20.9a"	"21.cie"
Feature 1: Allow catchability to vary?	no	yes	no	no	no
Feature 2: Allow domed survey selectivty?	no	no	yes	no	no
Feature 3: Use fishery CPUE?	no	no	no	yes	no
Feature 4: Estimate survey CV internally?	no	no	no	no	yes
Natural mortality	1	1	1	1	1
Growth	6	6	6	6	6
Ageing error	4	4	4	4	4
Stock-recruitment	2	2	2	2	2
Initial fishing mortality	1	1	1	1	1
Dirichlet-multinomial coefficients	3	3	3	3	3
Log catchability (survey)	1	1	1	1	1
Selectivity (fishery)	5	5	5	5	5
Selectivity (survey, ascending)	2	2	2	2	2
Selectivity (survey, top and descending)			3		
Log catchability (fishery)				1	
"Extra" survey standard deviation					1
True parameters	25	25	28	26	26
Total parameters	295	333	298	296	296

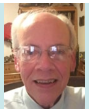


Results

Model names (1 of 2)

- Implications of data updates for names of existing models:
  - The SSC has stressed that model names should not change simply as a result of routine incrementing of existing time series (e.g., adding the most recent catch or survey index datum)
  - In keeping with the spirit of that policy, it seems that *any* sufficiently minor adjustments to existing time series should likewise not result in a new model name
  - Building upon the existing protocol, the criterion adopted here is that, *based on revisions to existing data alone*, a value of ADSB<0.05 does not merit a new model name

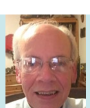
Update type	M19.12a	M19.12
Updated index data only	0.0243	0.0072
Updated index and agecomp data	0.0228	0.0095



Model names (2 of 2)

- Names of the new models:

Feature 1: Allow catchability to vary?	no	yes	no	no	no
Feature 2: Allow domed survey selectivty?	no	no	yes	no	no
Feature 3: Use fishery CPUE?	no	no	no	yes	no
Feature 4: Estimate survey CV internally?	no	no	no	no	yes
CIE review model name:	19.12a	19.12	"20.8a"	"20.9a"	"21.cie"
Average difference in spawning biomass:	n/a	n/a	0.4047	0.1299	0.1175
Final model name:	19.12a	19.12	21.1	21.2	21.3

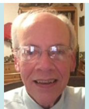


Objective function values

- Note that values are not strictly comparable across models

Model:	M19.12a	M19.12	M21.1	M21.2	M21.3
Allow catchability to vary?	no	yes	no	no	no
Allow domed survey selectivty?	no	no	yes	no	no
Use fishery CPUE?	no	no	no	yes	no
Estimate survey CV internally?	no	no	no	no	yes
Equilibrium catch:	0.00	0.00	0.00	0.00	0.00
Indices:	-3.87	-85.40	-4.15	21.64	-38.90
Sizecomps:	9335.21	9305.66	9321.96	9397.42	9291.17
Agecomps:	781.34	775.10	781.33	779.74	770.37
Recruitment:	-1.37	-1.75	-1.52	-1.83	-0.07
Initial recruitment:	5.08	6.38	3.25	5.05	6.35
Softbounds:	0.01	0.00	0.01	0.00	0.00
Parameter devs:	66.98	96.85	70.77	71.28	63.39
Total:	10183.38	10096.84	10171.64	10273.31	10092.32

NOAA FISHERIES

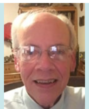


Fit to index data (1 of 4)

- Root-mean-squared-standardized-residual

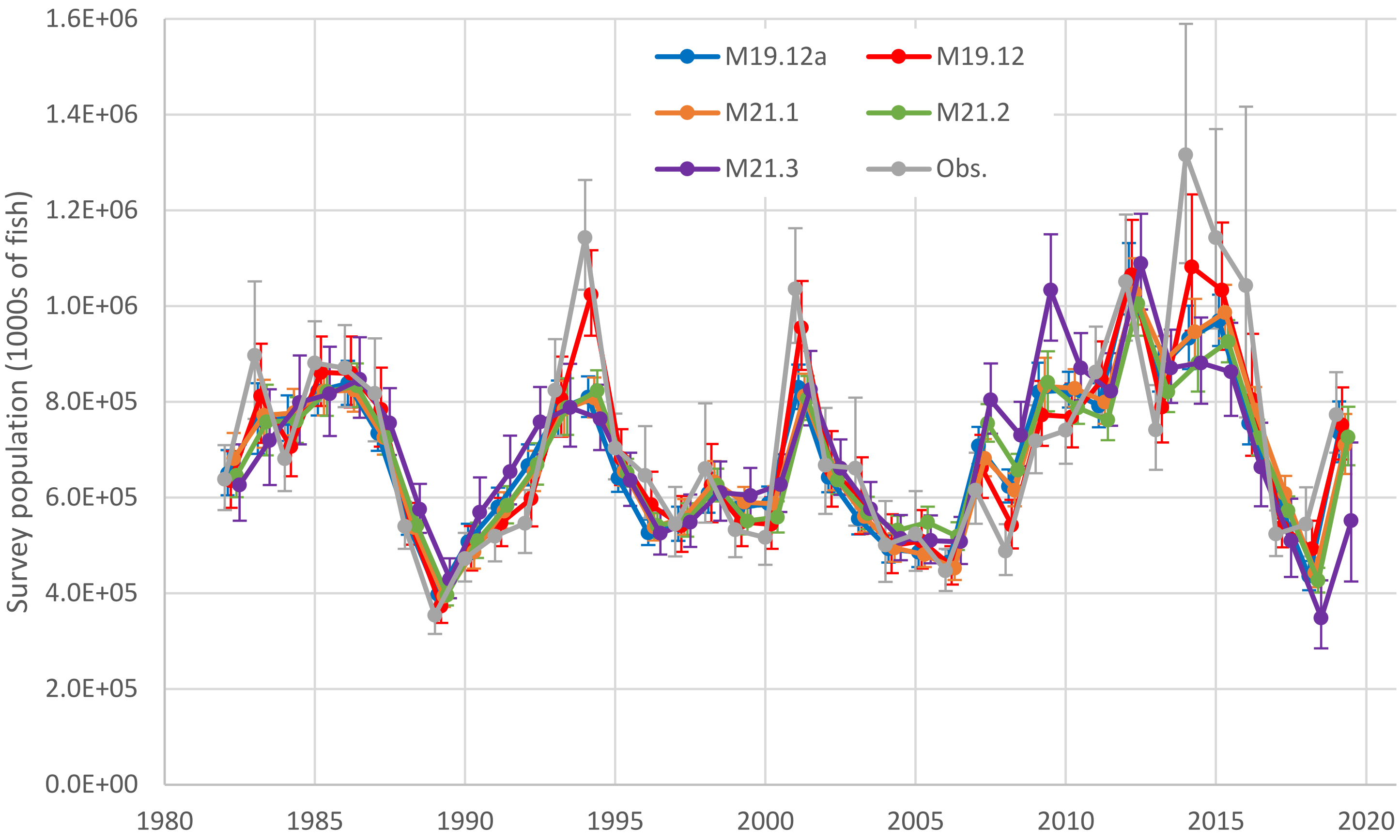
Index:	Survey					Fishery
Model:	M19.12a	M19.12	M21.1	M21.2	M21.3	M21.2
RMSSR:	2.301	1.002	2.298	2.425	1.002	2.561

- Only Models 19.12 and 21.3 meet the objective of RMSSR=1.0
  - Note that they do so by very different means, however



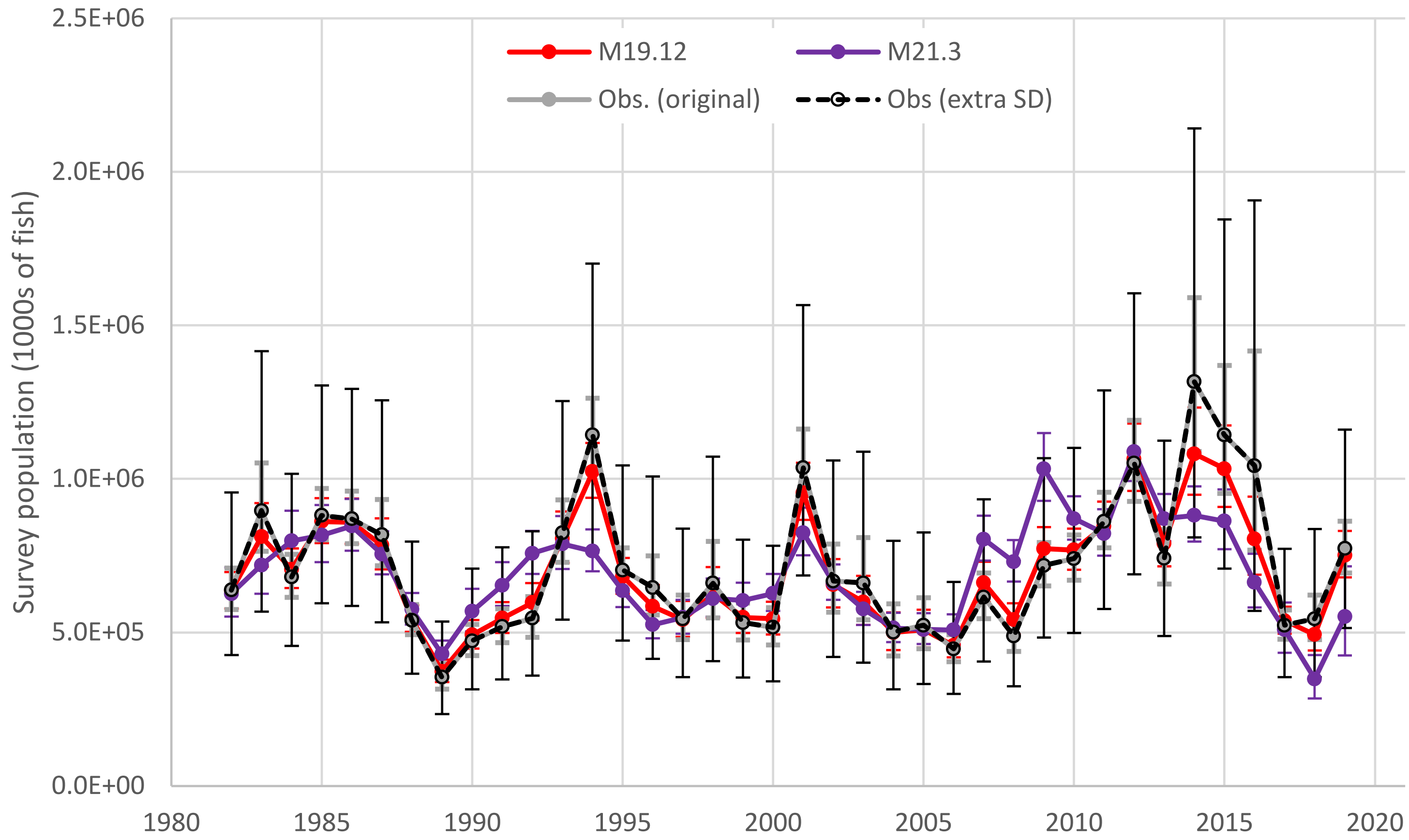
Fit to index data (2 of 4)

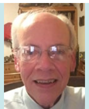
- Survey index, all models



Fit to index data (3 of 4)

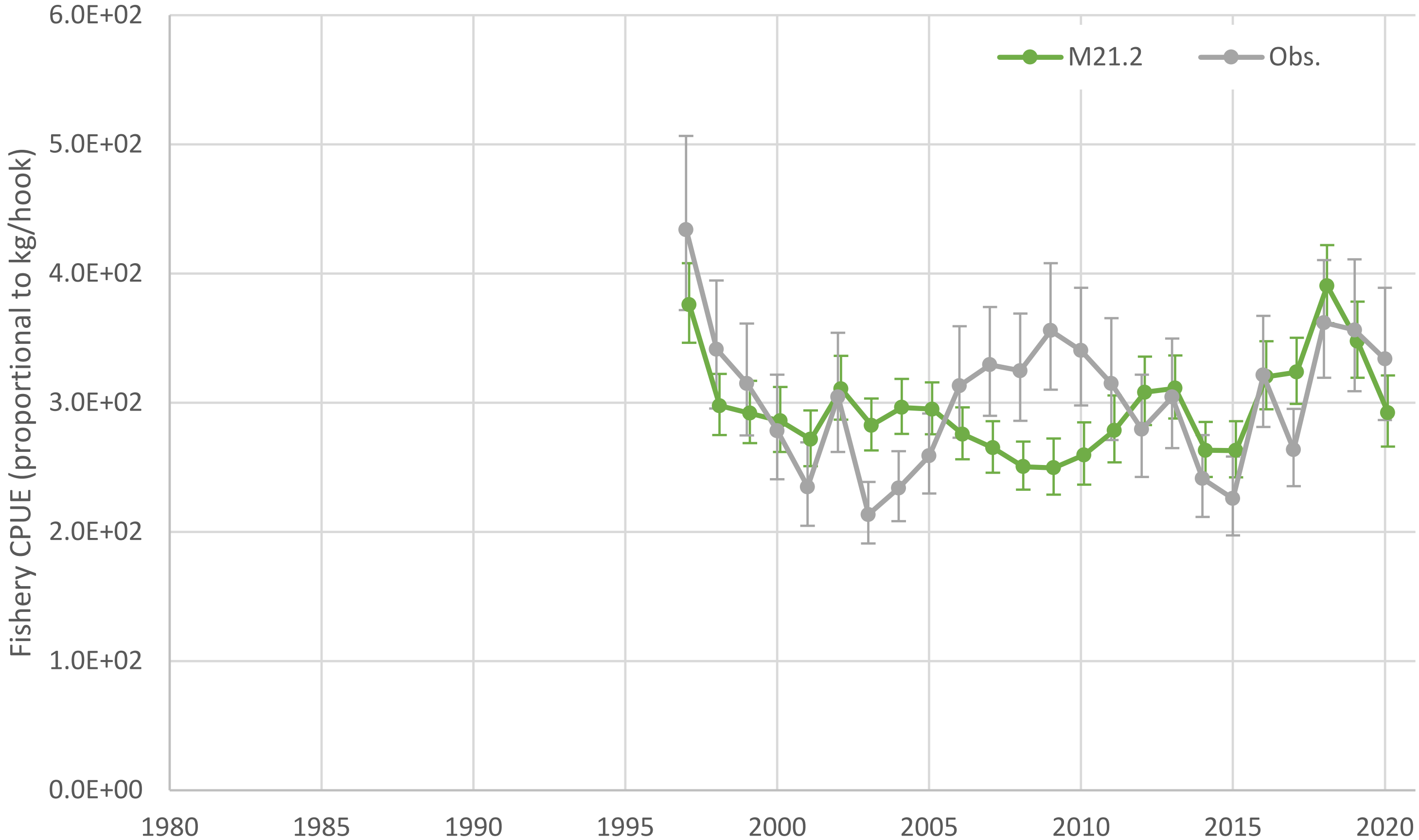
- Survey index (with and without “extra SD”), Models 19.12 and 21.3

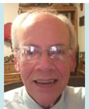




Fit to index data (4 of 4)

- Fishery CPUE index, Model 21.2 (correlation = 0.443)





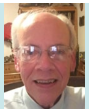
Fit to size composition data

- Fishery

	Model:	M19.12a	M19.12	M21.1	M21.2	M21.3
	Nave:	356	356	356	356	356
McAllister-Ianelli	Neff:	815	813	809	809	820
	Ratio:	2.292	2.286	2.275	2.275	2.305
Thorson et al.	$\ln(\theta)$:	10.000	10.000	10.000	9.989	10.000
	Neff:	356	356	356	356	356
	Ratio:	1.000	1.000	1.000	1.000	1.000

- Survey

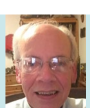
	Model:	M19.12a	M19.12	M21.1	M21.2	M21.3
	Nave:	356	356	356	356	356
McAllister-Ianelli	Neff:	596	621	603	570	636
	Ratio:	1.676	1.744	1.695	1.603	1.787
Thorson et al.	$\ln(\theta)$:	10.000	10.000	10.000	9.982	10.000
	Neff:	356	356	356	356	356
	Ratio:	1.000	1.000	1.000	1.000	1.000



Fit to age composition data

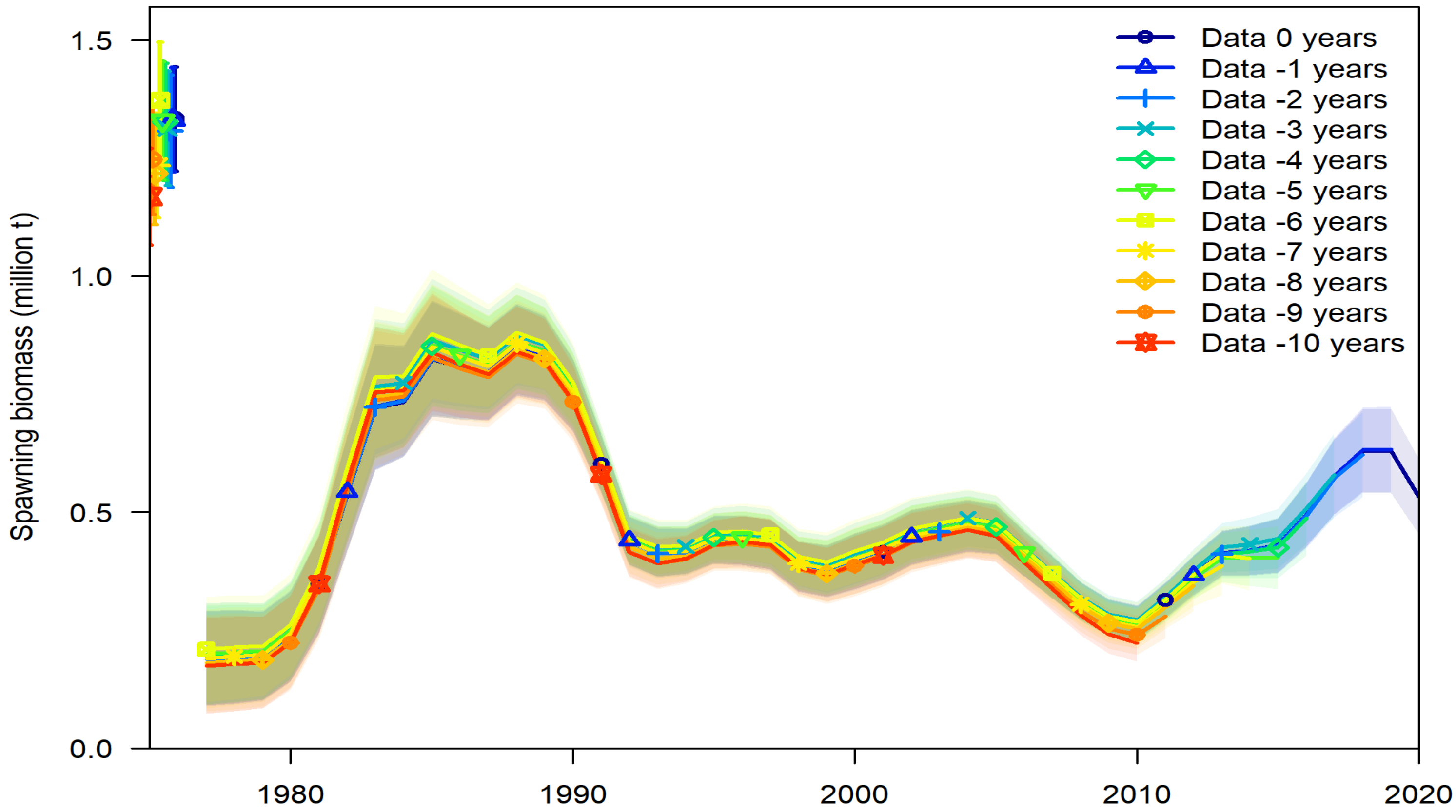
- Survey

	Model:	M19.12a	M19.12	M21.1	M21.2	M21.3
	Nave:	373	373	373	373	373
McAllister-Ianelli	Neff:	101	111	100	93	118
	Ratio:	0.272	0.299	0.268	0.250	0.316
Thorson et al.	$\ln(\theta)$:	-0.133	0.091	-0.291	-0.331	0.191
	Neff:	174	195	160	156	204
	Ratio:	0.468	0.524	0.429	0.419	0.549



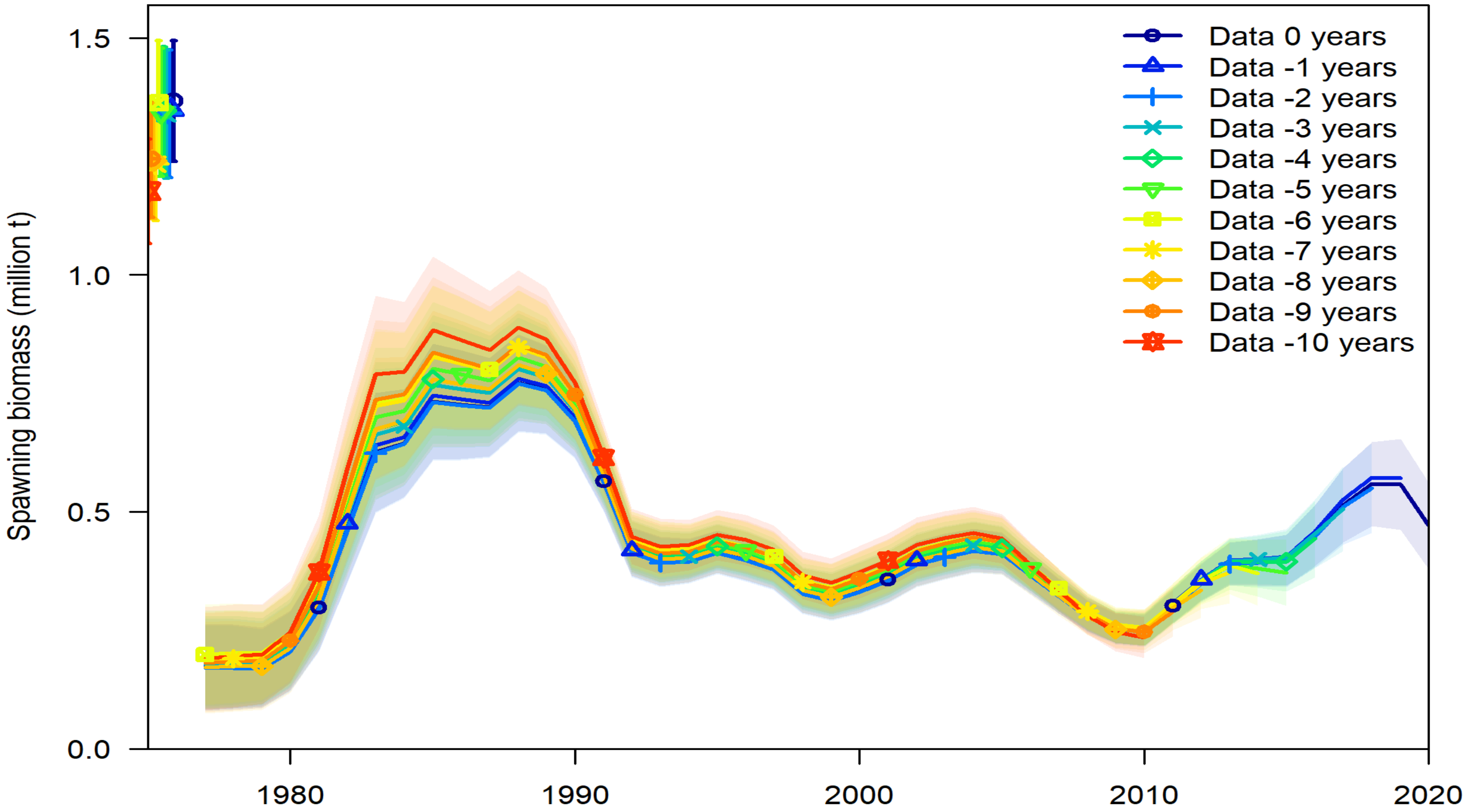
Retrospective behavior (1 of 5)

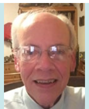
- Model 19.12a (ρ = -0.0500)



Retrospective behavior (2 of 5)

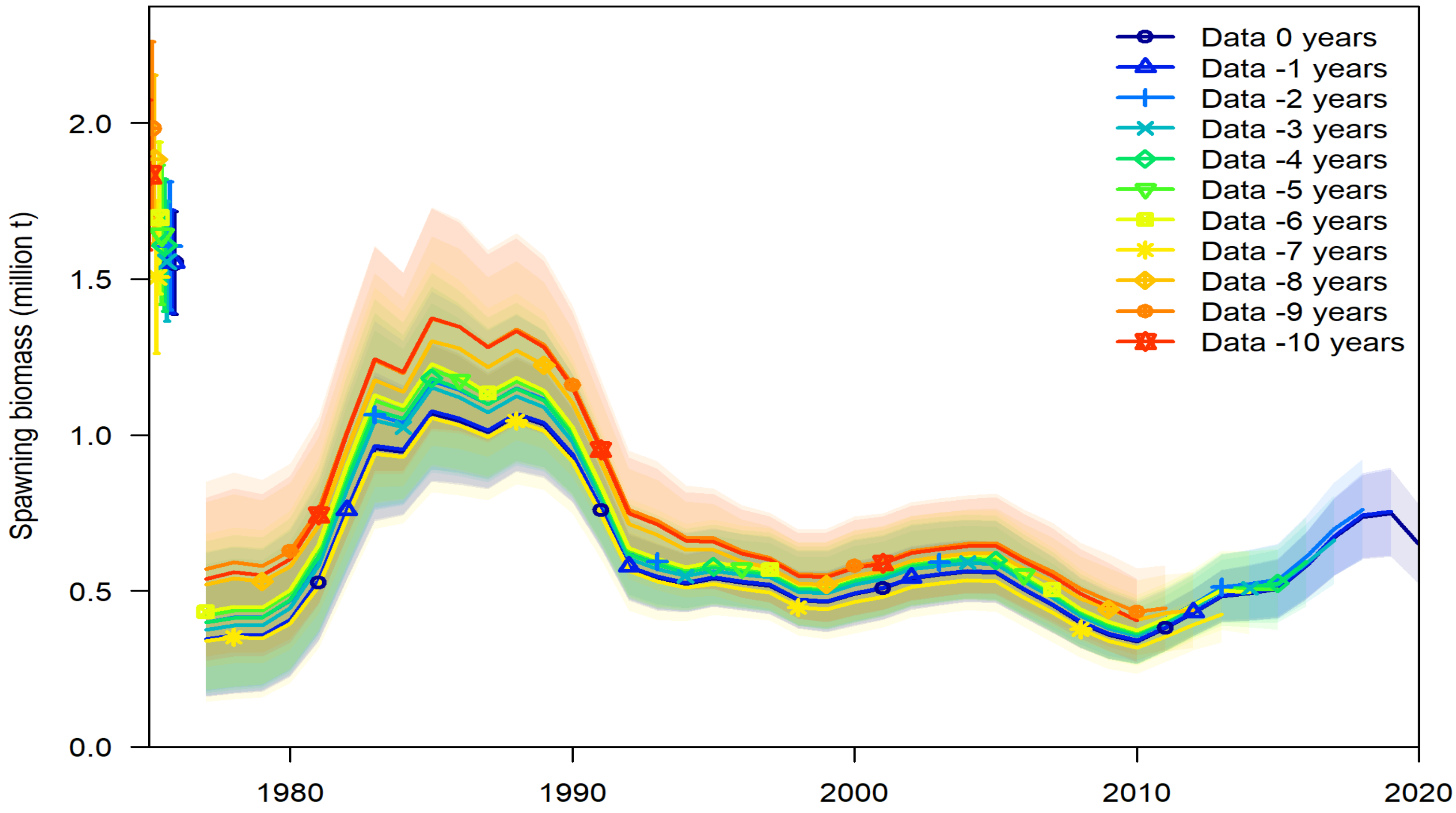
- Model 19.12 (ρ = -0.0352)





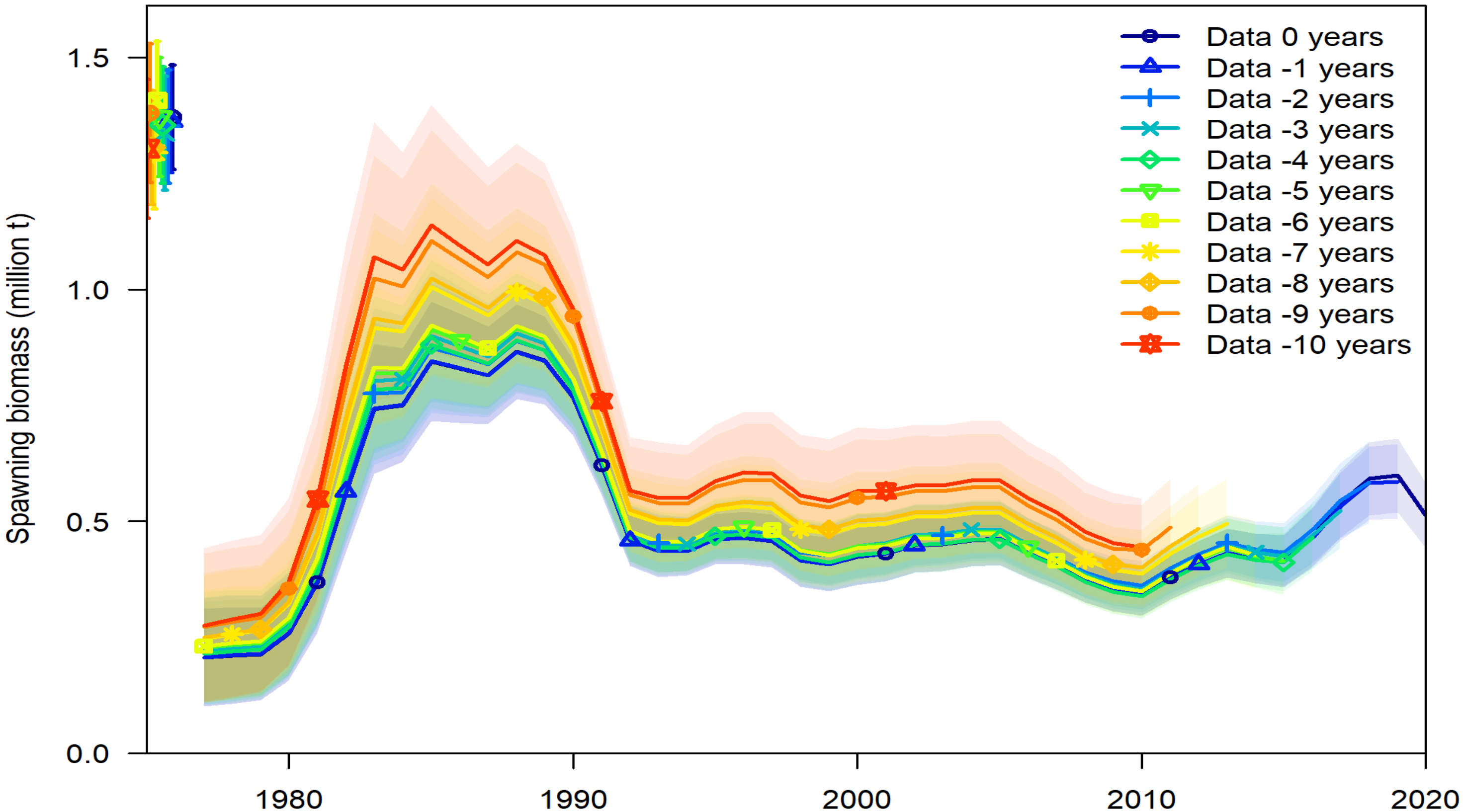
Retrospective behavior (3 of 5)

- Model 21.1 (ρ = 0.0326)



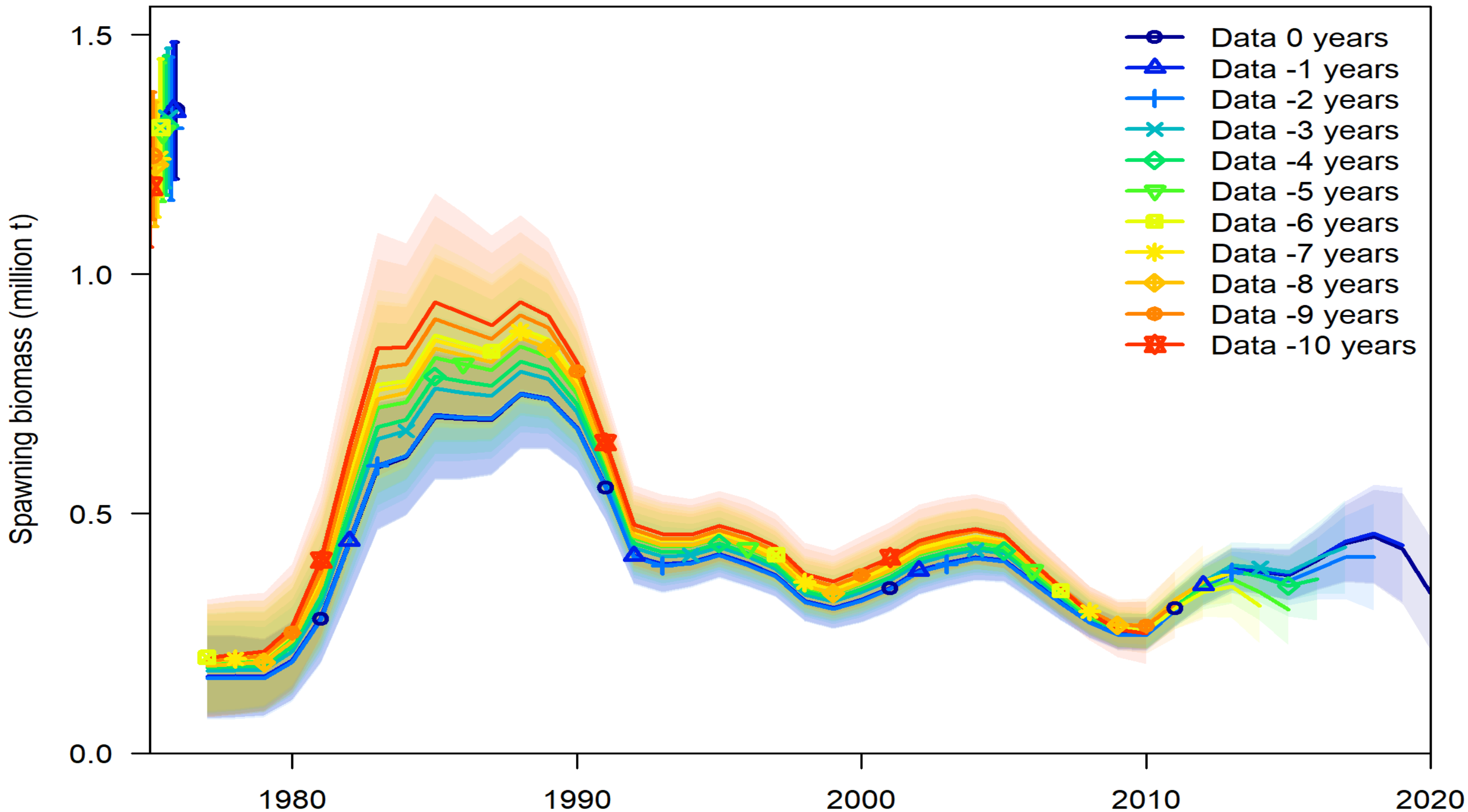
Retrospective behavior (4 of 5)

- Model 21.2 (ρ = 0.0875)



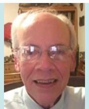
Retrospective behavior (5 of 5)

- Model 21.3 (ρ = -0.0535)



Time-invariant parameters

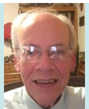
Model:	Model 19.12a		Model 19.12		Model 21.1		Model 21.2		Model 21.3	
Parameter	Est.	SD	Est.	SD	Est.	SD	Est.	SD	Est.	SD
Natural mortality	0.348	0.011	0.331	0.012	0.309	0.015	0.343	0.011	0.328	0.013
Mean length at age 1.5	14.777	0.387	14.877	0.392	14.819	0.374	14.645	0.391	15.004	0.399
Asymptotic length	112.948	3.052	120.718	5.047	103.223	2.670	116.248	3.482	122.377	5.471
Brody growth coefficient	0.118	0.009	0.099	0.011	0.158	0.013	0.104	0.009	0.097	0.011
Richards growth coefficient	1.439	0.042	1.498	0.047	1.287	0.052	1.517	0.042	1.491	0.048
SD(length at age 1)	3.485	0.067	3.501	0.066	3.507	0.069	3.506	0.069	3.476	0.065
SD(length at age 20)	9.905	0.380	10.071	0.463	9.062	0.365	10.167	0.432	10.182	0.485
Mean ageing bias at age 1	0.343	0.017	0.338	0.016	0.338	0.018	0.344	0.018	0.338	0.016
Mean ageing bias at age 20	1.116	0.226	1.195	0.221	1.214	0.241	1.100	0.236	1.205	0.218
Mean bias at age 1 (2008+)	0.002	0.025	0.007	0.024	0.004	0.026	0.005	0.026	0.005	0.025
Mean bias at age 20 (2008+)	-1.708	0.317	-1.924	0.317	-1.820	0.339	-1.882	0.343	-2.066	0.329
ln(mean post-1976 recruits)	13.129	0.096	12.979	0.099	12.940	0.117	13.130	0.098	12.921	0.106
ln(pre-1977 recruits offset)	-0.908	0.192	-0.945	0.182	-0.627	0.189	-0.880	0.188	-0.957	0.185
Pre-1977 fishing mortality	0.122	0.037	0.127	0.037	0.074	0.020	0.117	0.035	0.137	0.042
ln(Dirichlet-multinomial coef. for agecomps)	-0.133	0.192	0.091	0.221	-0.291	0.186	-0.331	0.176	0.191	0.240
ln(survey catchability)	0.003	0.062	0.099	0.064	0.094	0.078	-0.030	0.064	0.146	0.075
Fishery selectivity: begin flattop	74.984	0.039	74.867	0.515	72.179	0.718	75.949	0.061	74.990	0.519
Fishery selectivity: logit(flatop width)	-9.739	7.362	0.280	0.516	-9.669	9.085	-9.833	_	0.249	0.499
Fishery selectivity: ln(ascending SD)	5.914	0.028	5.909	0.038	5.853	0.042	5.968	0.031	5.911	0.037
Fishery selectivity: ln(descending SD)	-10.000	_	4.575	1.418	3.988	0.511	-8.275	14.966	4.595	1.341
Fishery selectivity: logit(ending value)	2.101	0.301	-2.828	3.088	0.765	0.333	1.856	0.271	-2.940	3.042
Survey selectivity: begin flattop	20.875	0.780	20.672	0.820	20.800	0.770	20.291	0.733	20.602	0.871
Survey selectivity: ln(ascending SD)	3.522	0.153	3.475	0.161	3.536	0.150	3.412	0.149	3.451	0.174
Survey selectivity: logit(flattop width)					-1.239	0.217				
Survey selectivity: ln(descending SD)					7.421	0.551				
Survey selectivity: logit(ending value)					-0.802	0.668				
ln(fishery catchability)							-5.952	0.064		
"Extra" survey standard deviation									0.152	0.030



Sigmas for deviation vectors

- Sigma for log catchability was tuned in Model 19.12 to a value of 0.0839
- Sigmas for other deviation vectors were tuned as follows:

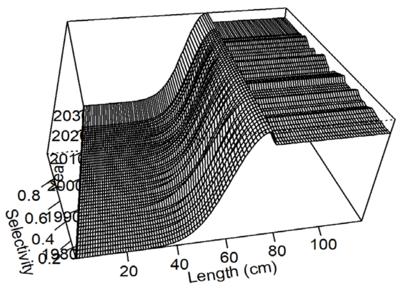
	Model 19.12a			Model 19.12			Model 21.1			Model 21.2			Model 21.3		
Parameter	var_dev	ave_var	sigma	var_dev	ave_var	sigma	var_dev	ave_var	sigma	var_dev	ave_var	sigma	var_dev	ave_var	sigma
ln(Recruits)	0.4434	0.0125	0.6765	0.4312	0.0129	0.6664	0.4407	0.0126	0.6733	0.4334	0.0131	0.6682	0.4547	0.0144	0.6852
Length_at_1.5	0.7944	0.1985	0.1474	0.7991	0.2019	0.1486	0.8040	0.1967	0.1422	0.7951	0.2060	0.1480	0.7950	0.1985	0.1504
Sel_fsh_lnSD1	0.7089	0.2897	0.1559	0.7097	0.2913	0.1542	0.7112	0.2883	0.1734	0.7466	0.2496	0.1721	0.6994	0.3017	0.1482
Sel_fsh_logitEnd	0.1913	0.8072	0.7525	0.0002	0.9989	0.7640	0.3612	0.6457	0.6288	0.4351	0.5581	1.1255	0.0002	0.9989	0.7350
Sel_srv_PeakStart	0.8508	0.1589	0.2035	0.8552	0.1480	0.2204	0.8466	0.1569	0.2011	0.8208	0.1815	0.1875	0.8582	0.1395	0.2380
Sel_srv_lnSD1	0.7332	0.2752	0.7691	0.7521	0.2539	0.8365	0.7267	0.2789	0.7519	0.6631	0.3410	0.6889	0.7683	0.2272	0.9320



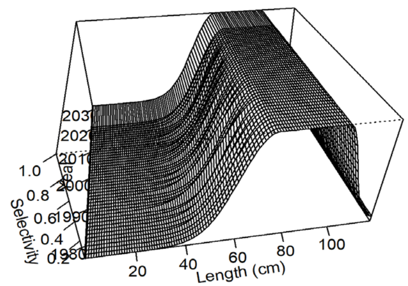
Selectivity

Fishery

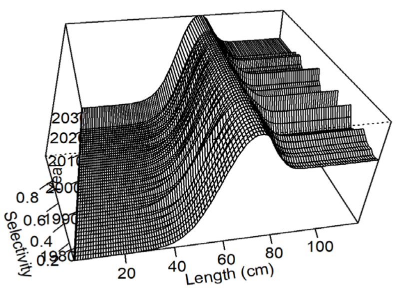
Model 19.12a



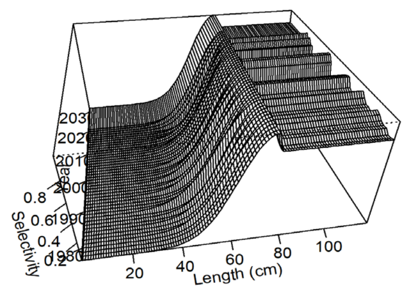
Model 19.12



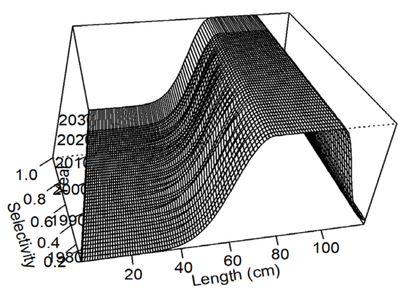
Model 21.1



Model 21.2

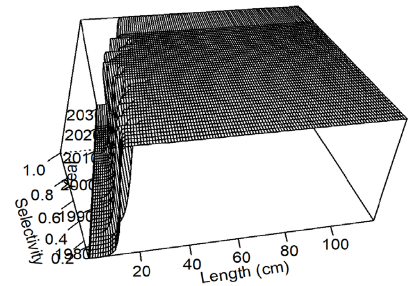


Model 21.3

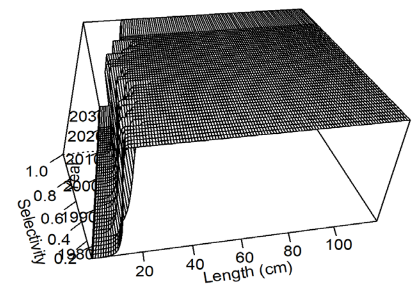


Survey

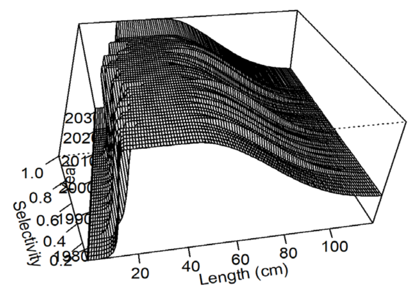
Model 19.12a



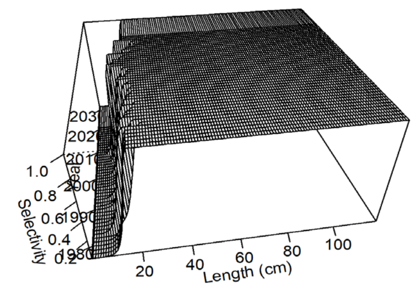
Model 19.12



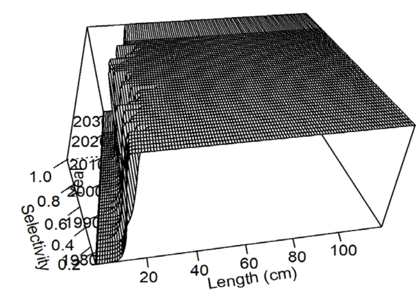
Model 21.1

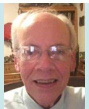


Model 21.2



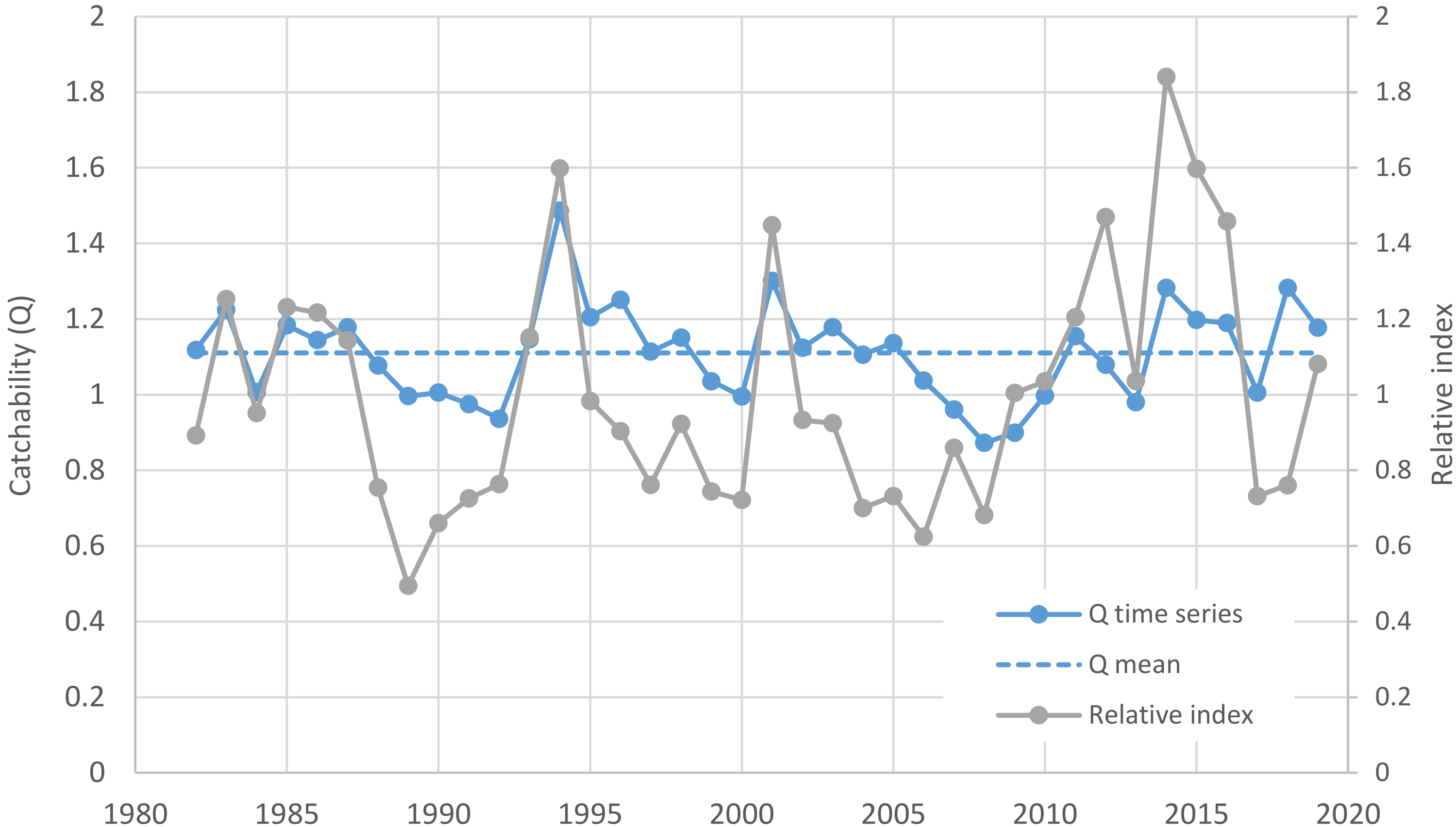
Model 21.3





Time-varying survey catchability (Model 19.12)

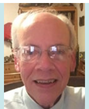
- Compared with relative survey index (correlation = 0.622)



Model weighting (1 of 2)

- An ensemble consists of:
  1. a set of models, and
  2. a set of weights
- The last two assessments, with Team and SSC approval, computed model weights as an emphasis-weighted average of (0,1) scores for each member of a set of ranking criteria
- CIE also adopted this approach, with some modifications:
  - Some new ranking criteria added and some old ones removed
  - Instead of (0,1) scores, (0,1,2) scores averaged across reviewers
  - For any criteria with equal average scores across models:
    - Emphasis set at 0, to avoid skewing toward equal weighting
    - Kept in the table nevertheless (for potential use in the future)

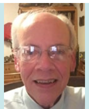
NOAA FISHERIES



Model weighting (2 of 2)

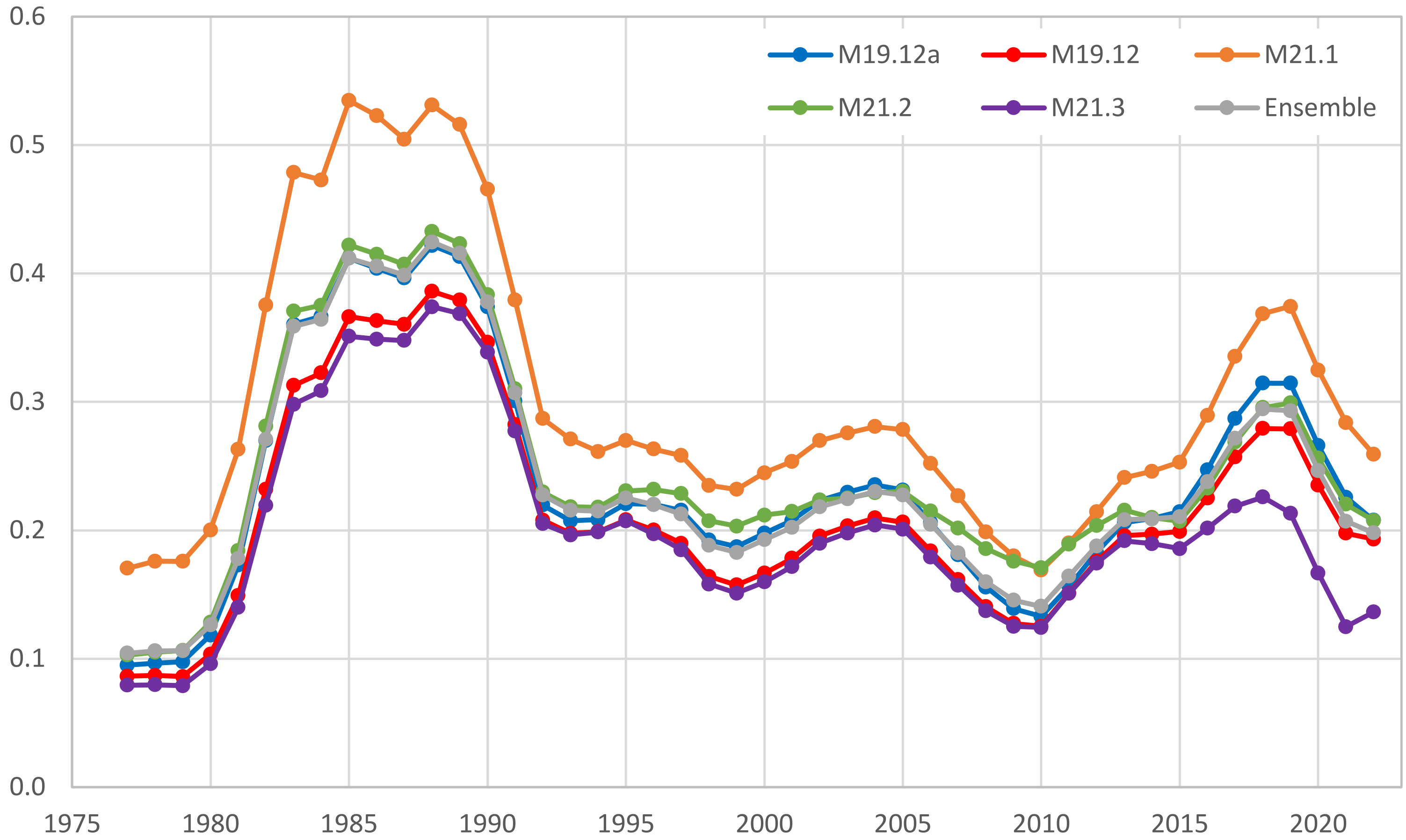
- CIE conclusions (which the reviewers anticipated would be revised)

Feature/criterion/result	Emph.	19.12a	19.12	21.1	21.2	21.3
Feature 1: Allow catchability to vary?		no	yes	no	no	no
Feature 2: Allow domed survey selectivity?		no	no	yes	no	no
Feature 3: Use fishery CPUE?		no	no	no	yes	no
Feature 4: Estimate survey CV internally?		no	no	no	no	yes
General plausibility of the model	3	2	1	0.6667	1	1.3333
Acceptable retrospective bias	3	2	2	1.3333	1	2
Uses properly vetted data	3	2	2	2	0	2
Acceptable residual patterns	3	2	2	2	2	1
Comparable complexity	2	2	1	1	2	2
Fits consistent with variances	2	1	2	1	0	2
Dev sigmas estimated appropriately	0					
Incremental changes	0					
Objective criterion for sample sizes	0					
Change in ageing criteria addressed	0					
Density dependence (other than R) addressed	0					
Regime shifts addressed	0					
Average emphasis:		0.9375	0.8438	0.6875	0.5000	0.8438
Model weight:		0.2459	0.2213	0.1803	0.1311	0.2213



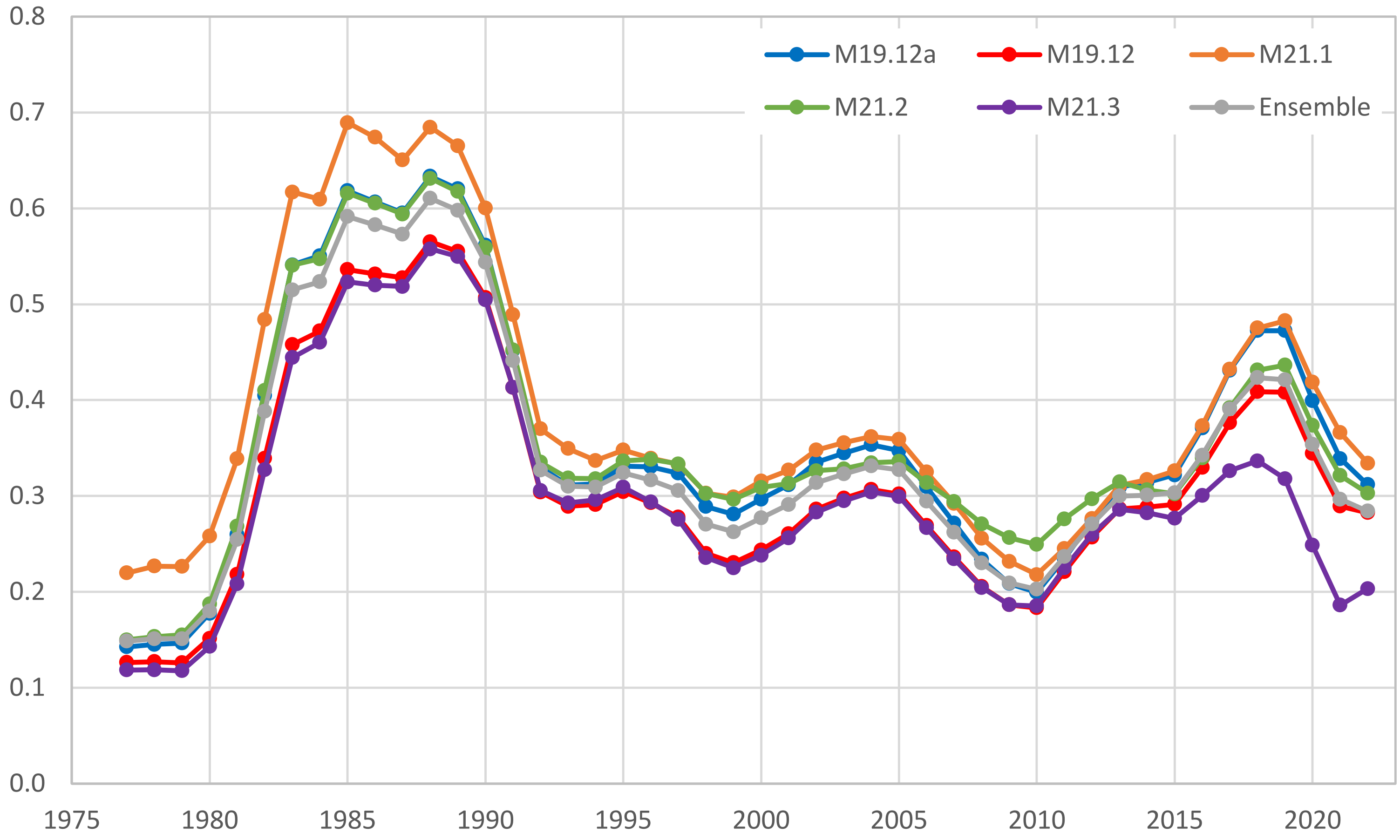
Time series: female spawning biomass

- Values are in millions of t



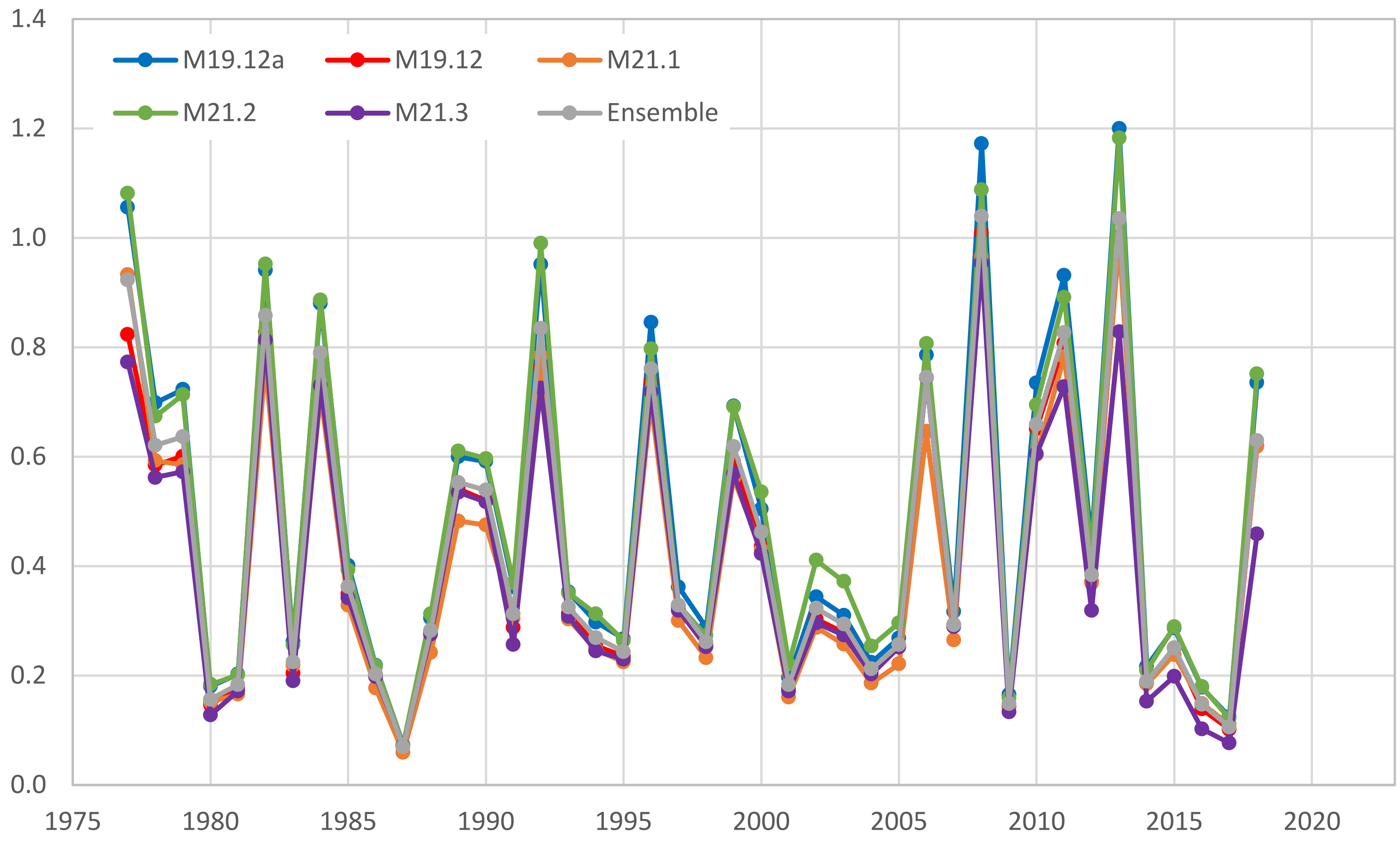
Time series: relative spawning biomass

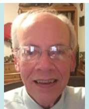
- Values are relative to $B_{100\%}$



Time series: age 0 recruitment

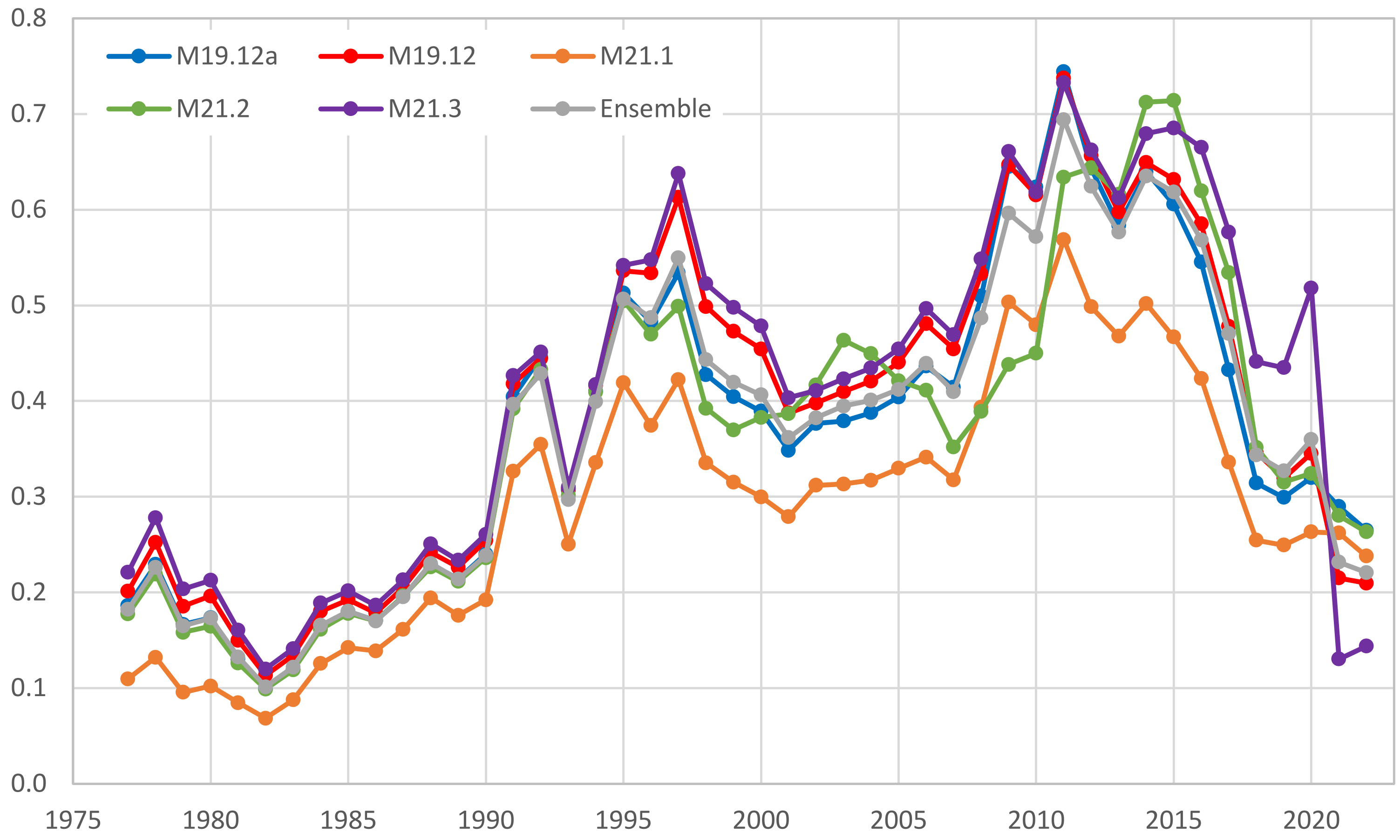
- Values are in billions of fish

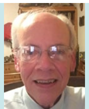




Time series: fishing mortality rate

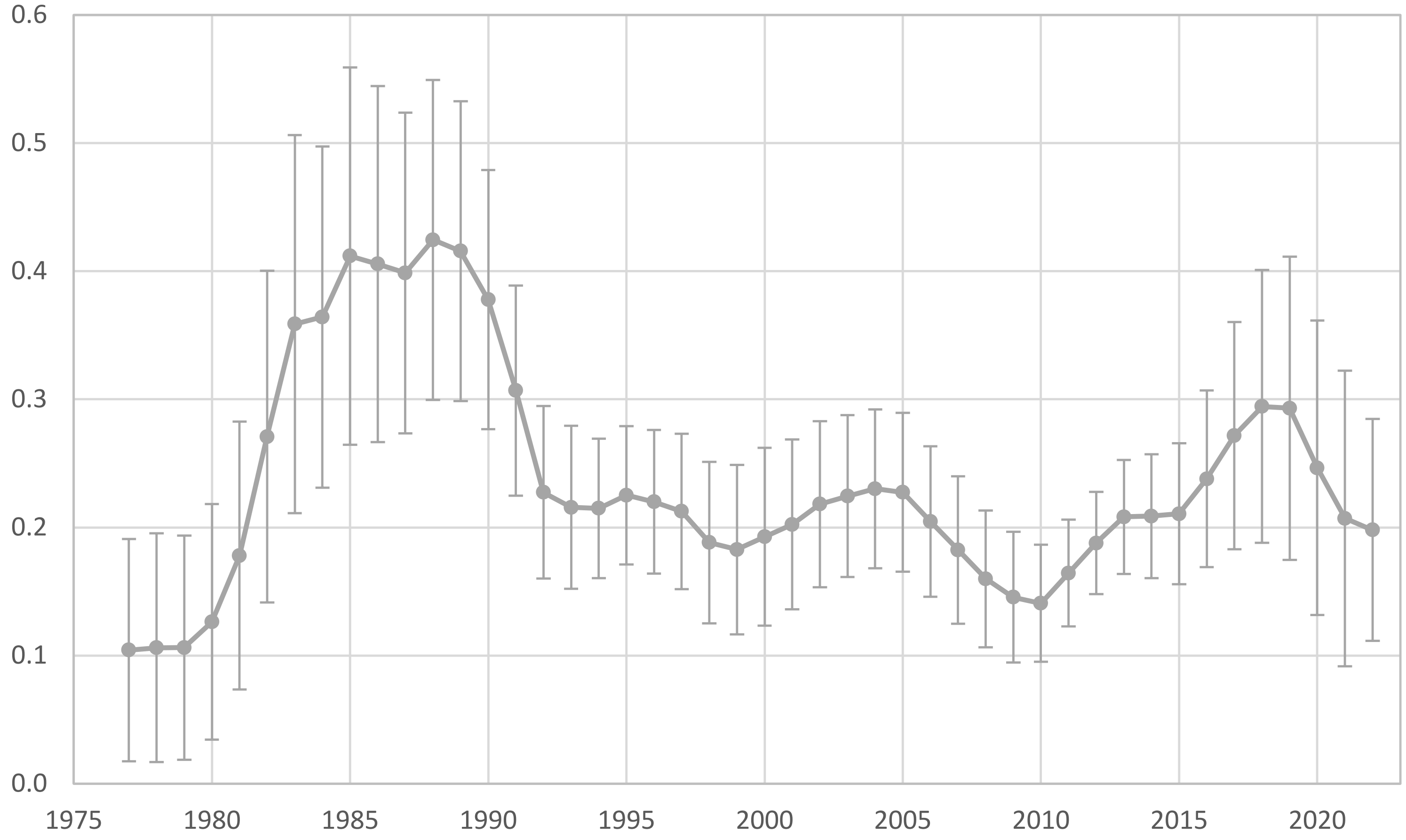
- 2021-2022 values are *not* conditioned on the specified 2021 ABC





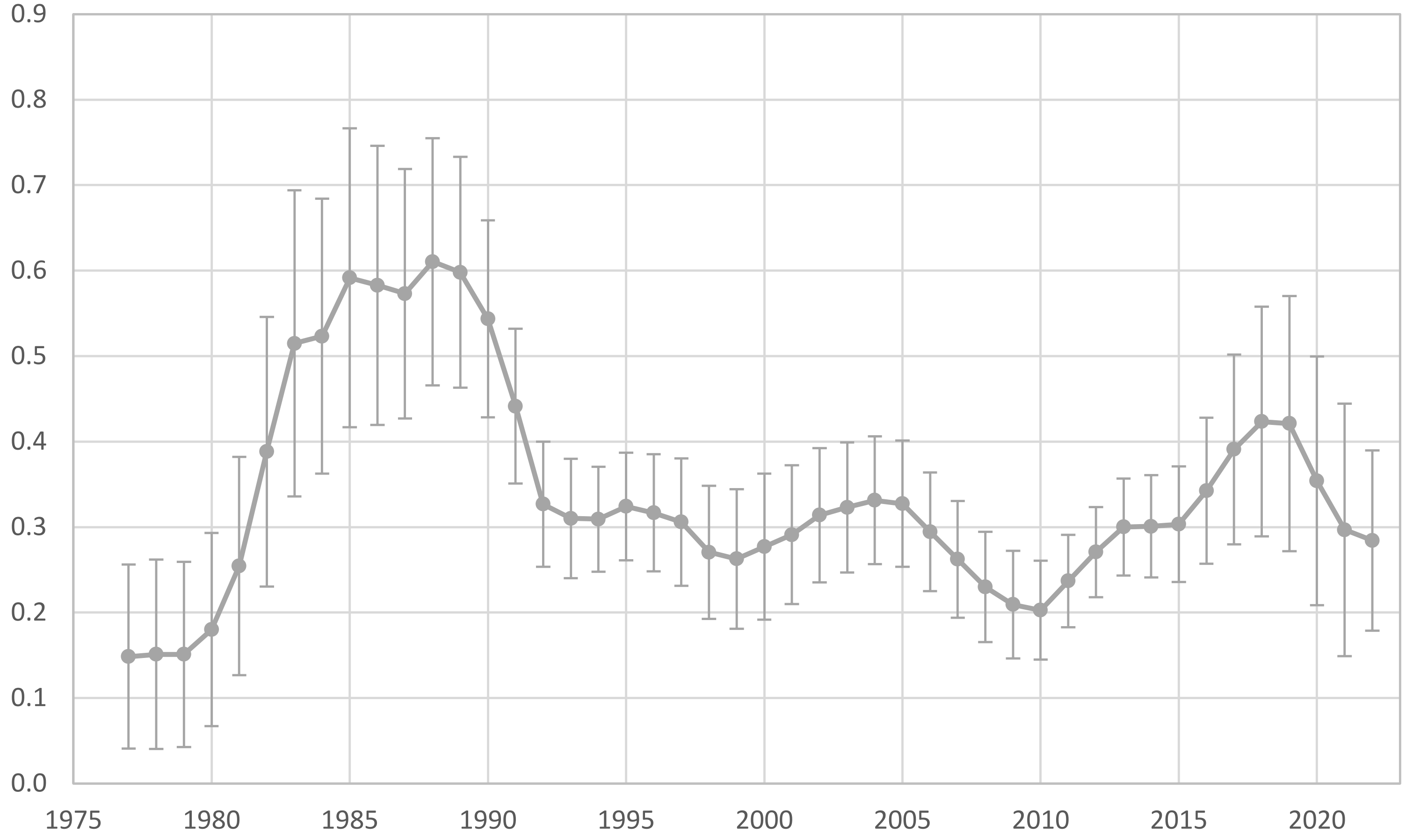
Ensemble: female spawning biomass

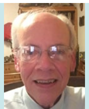
- Error bars = +/- 2 standard deviations; values are in millions of t



Ensemble: relative spawning biomass

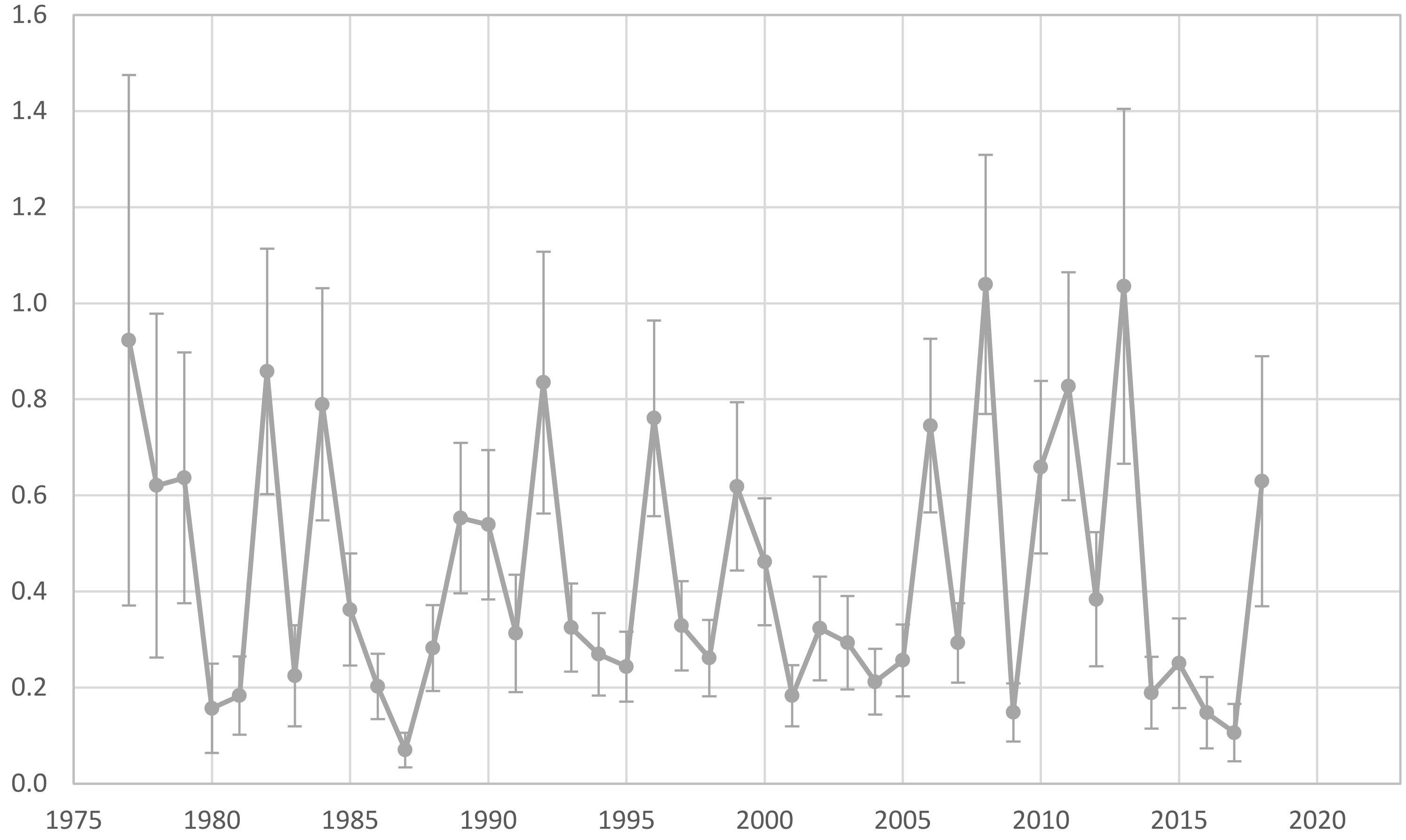
- Error bars = +/- 2 standard deviations; values are relative to $B_{100\%}$

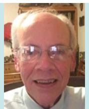




Ensemble: age 0 recruitment

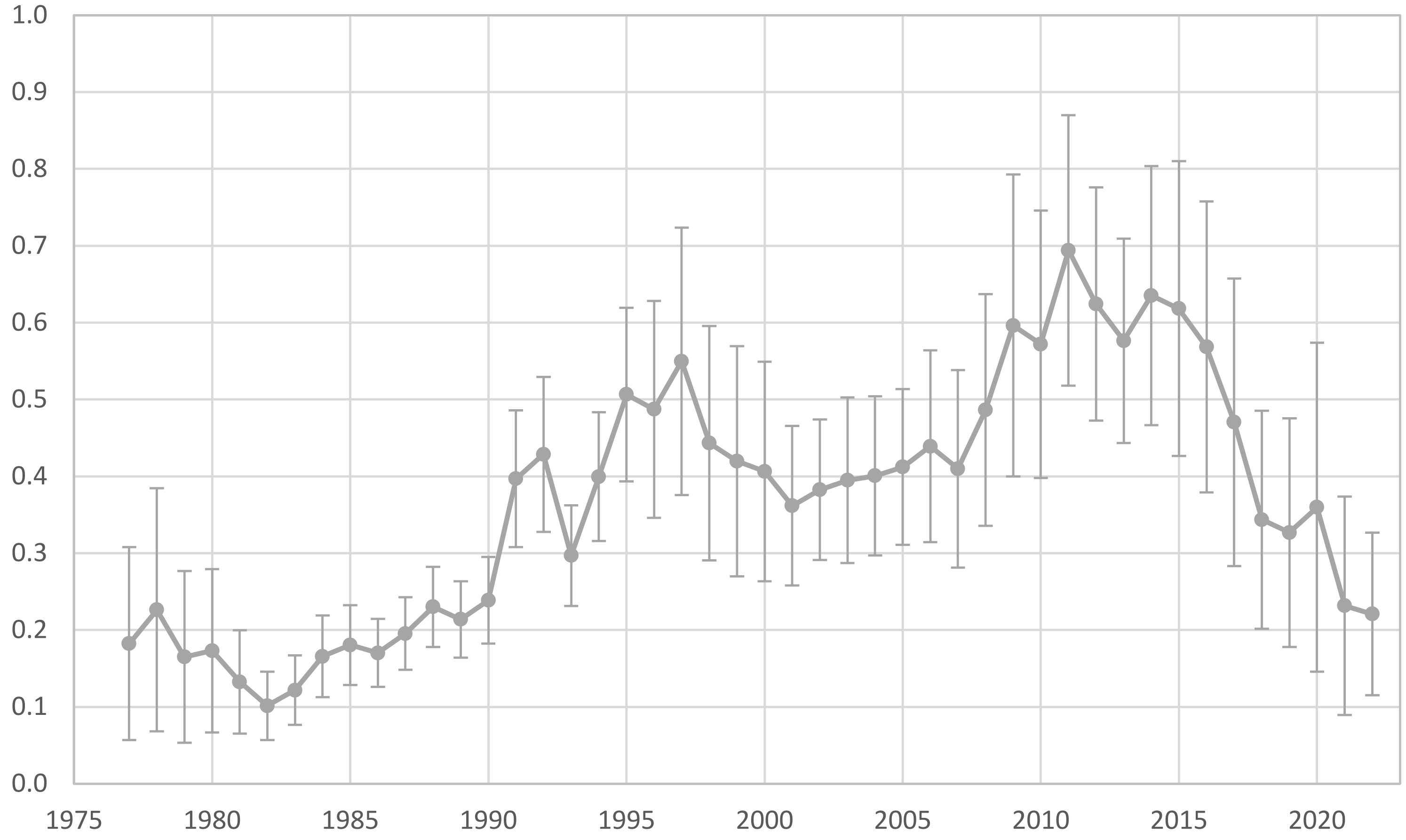
- Error bars = +/- 2 standard deviations; values are in billions of fish

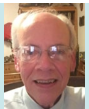




Ensemble: fishing mortality rate

- 2021-2022 values are *not* conditioned on the specified 2021 ABC



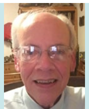


2021-2022 ABC and OFL summary statistics

- Specified 2021 values: ABC = 123805, OFL = 147949
- Specified 2022 values: ABC = 106852, OFL = 128340

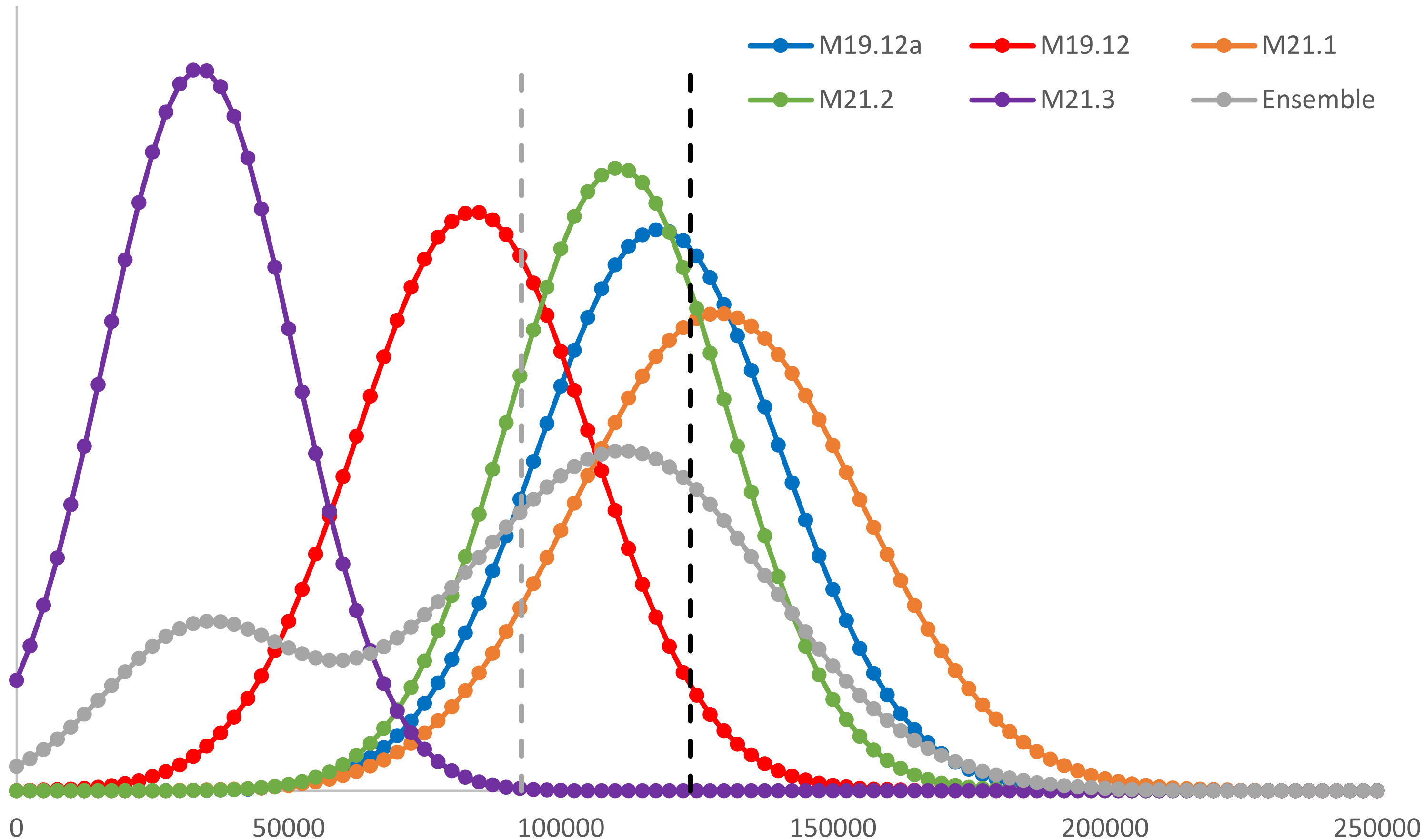
Feature			19.12a	19.12	21.1	21.2	21.3	
Allow catchability to vary?			no	yes	no	no	no	
Allow domed survey selectivity?			no	no	yes	no	no	
Use fishery CPUE?			no	no	no	yes	no	
Estimate survey CV internally?			no	no	no	no	yes	
Model weight:			0.2459	0.2213	0.1803	0.1311	0.2213	
Year	Quantity	Statistic	19.12a	19.12	21.1	21.2	21.3	Ensemble
2021	ABC	mean	118044	83930	128897	110619	33599	92789
2021	ABC	sdev	22316	21641	26245	20106	17347	41186
2021	OFL	mean	141089	100683	152279	132399	40597	110785
2021	OFL	sdev	26414	25811	30727	23883	20857	48690
2022	ABC	mean	105613	82924	115920	102594	41566	87880
2022	ABC	sdev	12059	14233	15254	11607	15769	30379
2022	OFL	mean	117275	93561	128750	114476	48335	98472
2022	OFL	sdev	12304	15085	15880	11993	17699	32692

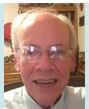
NOAA FISHERIES



ABC and OFL distributions: 2021 ABC

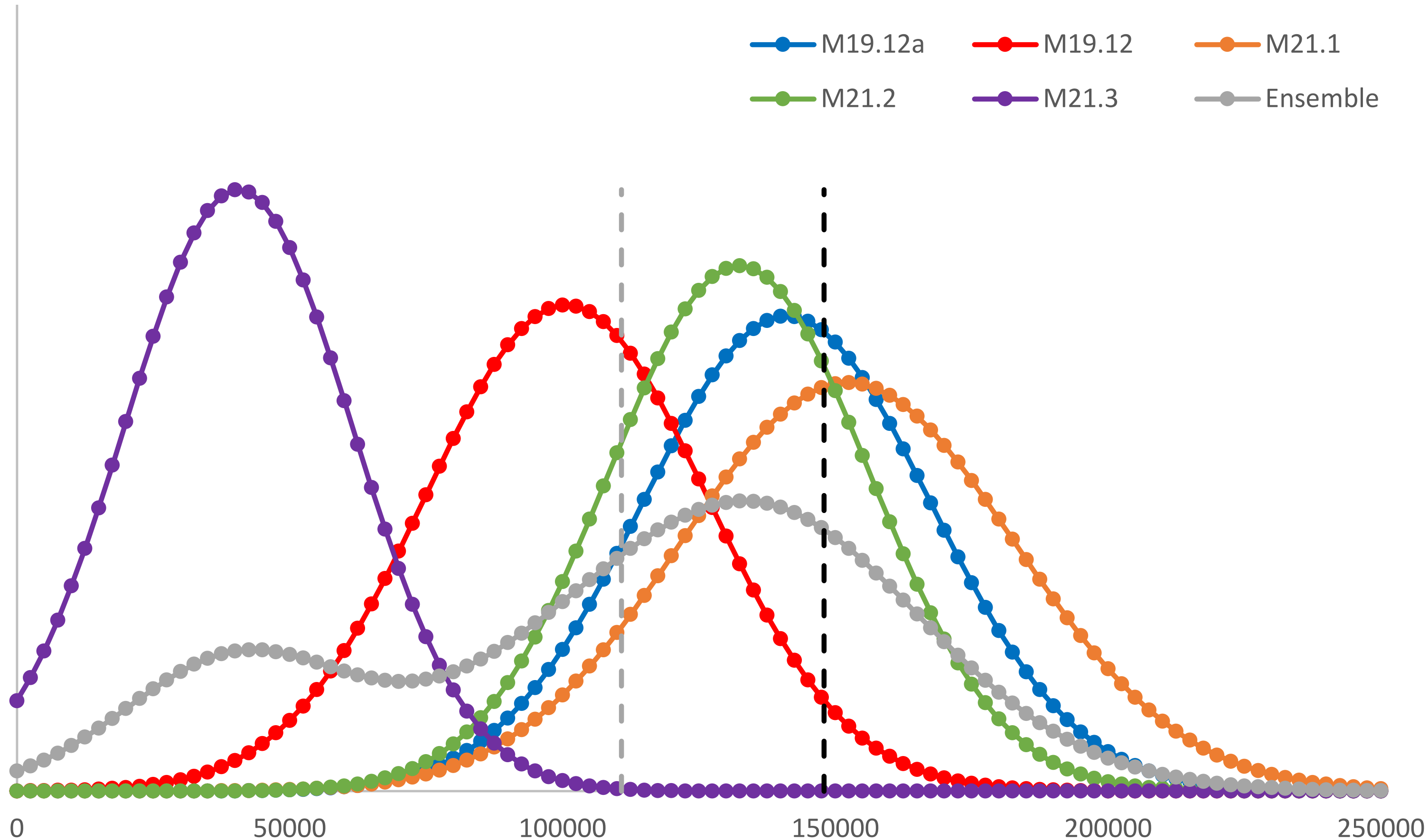
- Ensemble mean is 25% lower than the specified value





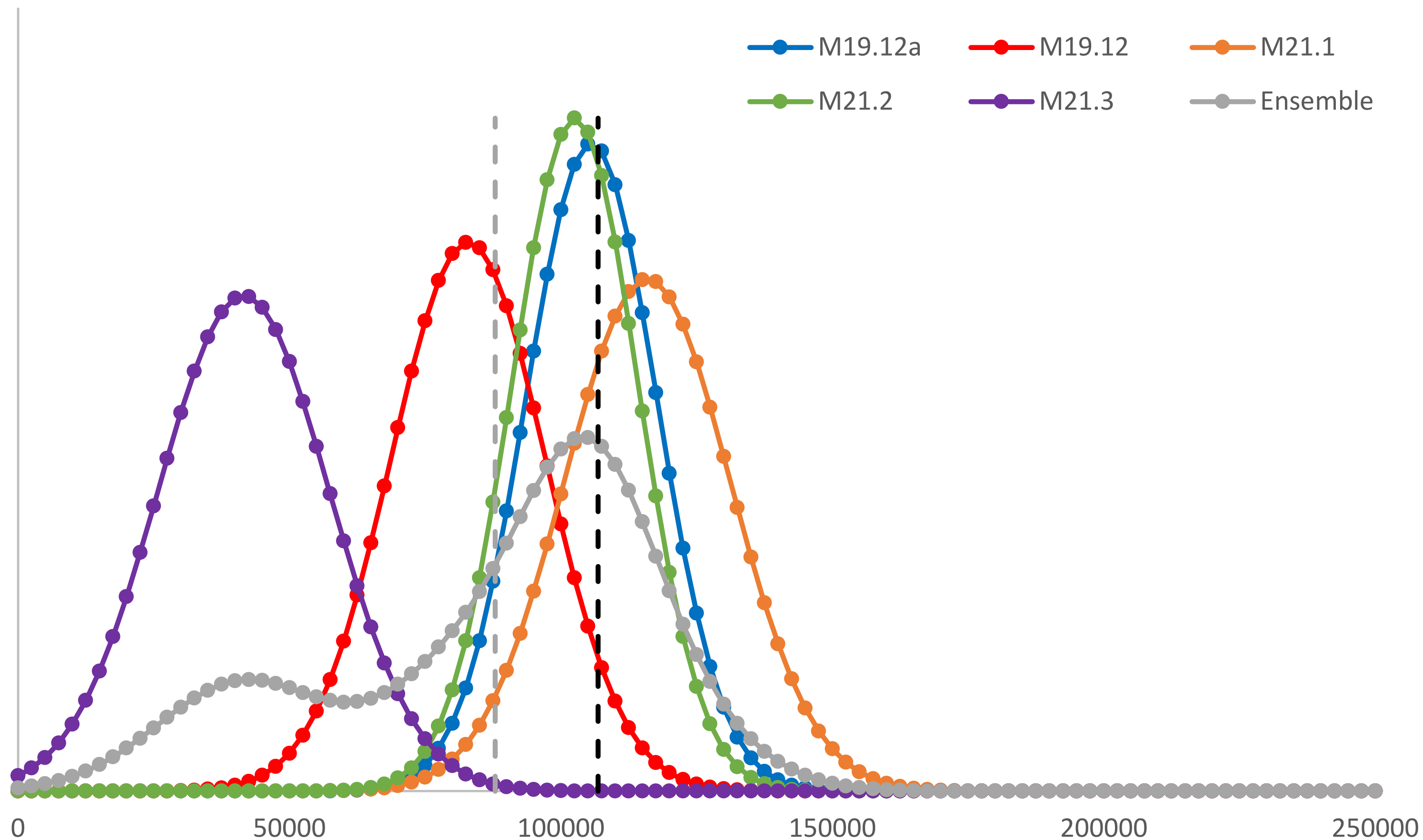
ABC and OFL distributions: 2021 OFL

- Ensemble mean is 25% lower than the specified value



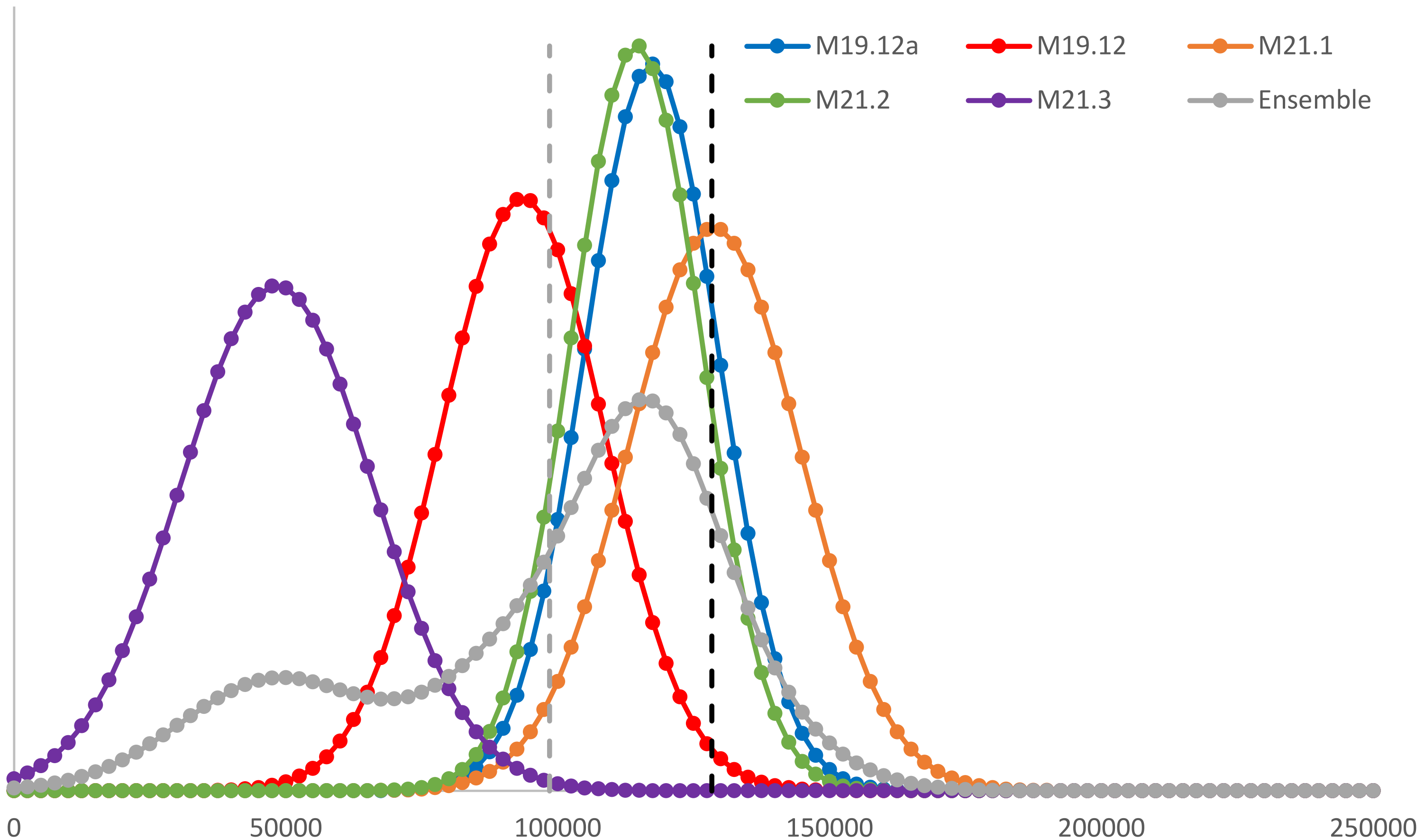
ABC and OFL distributions: 2022 ABC

- Ensemble mean is 18% lower than the specified value



ABC and OFL distributions: 2022 OFL

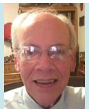
- Ensemble mean is 23% lower than the specified value



Discussion

Ensemble evaluation: Introduction

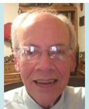
- Although this preliminary assessment is based on the ensemble (both the set of models and their respective weights) recommended by the CIE reviewers, the authors recognize that the Team, the SSC, or the authors themselves may recommend use of a different ensemble, or no ensemble at all, in the final assessment
- It should also be emphasized that the CIE reviewers anticipated that the Team or SSC would provide their own scores for the model weighting criteria, thus resulting a revised set of model weights
- Comments on individual models follow on the next five slides



Ensemble evaluation: Model 19.12a

- This is the base model for the current assessment, having been adopted by the SSC at the conclusion of the 2020 assessment cycle
- In many respects, it performs very well
- The CIE reviewers gave it (unanimously) the highest possible score for all but one of the ranking criteria, and it was explicitly endorsed as the single best model in the ensemble by at least two of the CIE reviewers
- Model 19.12a is also the most parsimonious model in the ensemble
- However, while this model clearly tracks the survey index to an appreciable degree (correlation = 0.853), the fit to those data is less than fully satisfactory, statistically speaking (RMSSR = 2.301)

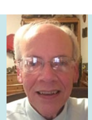
NOAA FISHERIES



Ensemble evaluation: Model 19.12

- Replaced as the base model at the conclusion of the 2020 cycle
- Time-varying catchability has been addressed several times, but has always been controversial, with 2019 being the only year in which a model with this feature was adopted by the SSC
- Arguments against: danger of over-parameterization, lack of an identified mechanism, decreased impact of the survey index on model results
- Argument in favor: needed in order to achieve a fit to the index data consistent with the uncertainty (as estimated outside the model)
- Wilberg et al. (2010): time-varying catchability should be the “default assumption,” particularly if the survey does not cover the stock’s range
- O’Leary et al. (2021): stock-wide (EBS, NBS, WBS) survey data show that “availability” in the 2017 EBS survey was 27% lower than in 2010

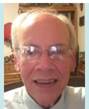
NOAA FISHERIES



Ensemble evaluation: Model 21.1

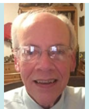
- Dome-shaped survey selectivity was a regular feature prior to 2016
- Team (9/15): “Dome-shaped survey selectivity seems inescapable”
- SSC (10/16): “In spite of the concerns over dome-shaped survey selectivity…, there are many potential mechanisms”
- Weinberg et al. (2016): “The results of our experiment do not support the use of a dome-shaped survey selectivity function”
- Recommendation following the 2016 CIE review: use “the simplest selectivity form that gives a reasonable fit;” logistic met this standard
- As prior to 2016, allowing for dome-shaped survey selectivity in this assessment resulted in a pronounced decrease at larger sizes
- This had a substantial effect on estimates of quantities such as spawning biomass, fishing mortality, and ABC and OFL; but little improvement in goodness of fit relative to Model 19.12a

NOAA FISHERIES



Ensemble evaluation: Model 21.2

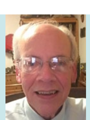
- Use of fishery CPUE data as an index of abundance has long been associated with a number of concerns
- The new fishery CPUE index was originally intended simply as the first step in what was anticipated to be a multi-year process of development
- Although the CIE reviewers assigned Model 21.2 a score of 0 under the "Uses properly vetted data" criterion, they nevertheless recommended including it in the ensemble, rather than waiting for further development
- One factor complicating the use of the new index in the model is that the index is specific to the longline fishery in the Jan-Feb period, whereas the model is aggregated across gear types and seasons
- In terms of spawning biomass, ABC, and OFL, Model 21.2 is the closest of the alternative models in the ensemble to the base model
- In terms of goodness of fit, Model 21.2 generally performs slightly less well than the base model, because it has to fit the fishery CPUE index



Ensemble evaluation: Model 21.3

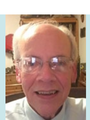
- Another model incorporating the “extra SD” feature was considered, but not accepted, during the 2017 assessment cycle (Model 17.3)
- Summing the estimated “extra SD” term with the log-scale standard errors gives values that are more than triple the original, on average
- Model 21.3 estimates that spawning biomass is currently below the $B_{20\%}$ threshold that results in closure of the directed fishery
- Like Model 19.12, the fit to the survey index is very good, but the mechanisms by which those two models achieve that result are different
- Similar to Model 19.12, one argument against use of Model 21.3 might be the decreased impact of the survey index on model results
- The CIE reviewers noted the recent string of positive residuals in the fit to the survey index, giving it a low “Acceptable residual patterns” score
- Wilberg et al. (2010) noted that inflating the standard deviations will often produce trends in residuals if catchability is actually time-varying

NOAA FISHERIES



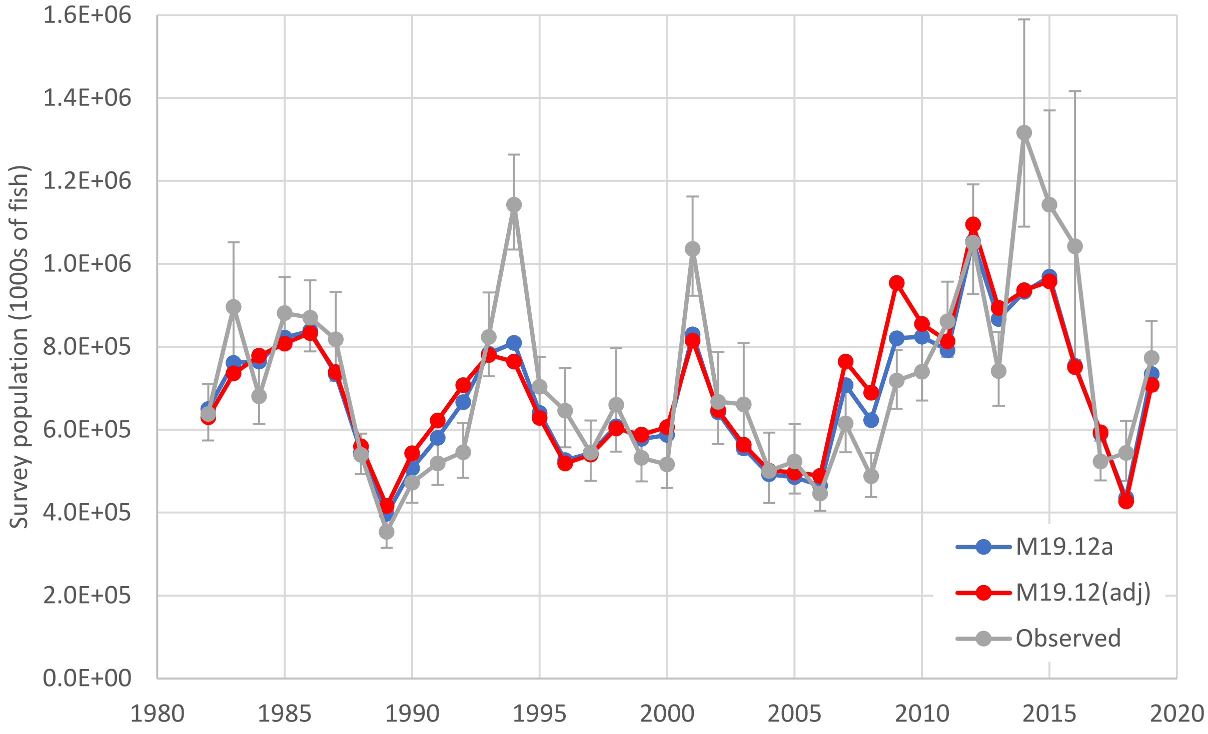
Thoughts on fitting the survey index (1 of 2)

- Two views of the standard errors estimated by VAST:
  1. They are too small
     - Solution: Estimate the “underage” and fit the data (Model 21.3)
  2. They are accurate
     - Solutions:
       - Parameterize so as to fit the data (Model 19.12)
       - Accept a substantial lack of fit (Model 19.12a)
- “Down-weighting” or “right-weighting?”
  - Is the goal to *maximize* the impact of the survey index, or to weight it appropriately?
  - Any time a model includes data in addition to the survey index, the impact of the survey index is inherently reduced

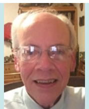


Thoughts on fitting the survey index (2 of 2)

- Comparing fits of M19.12a and M19.12 with time-varying Q removed



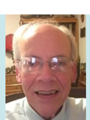
- RMSSR: 2.892 for M19.12(adj) versus 2.301 for M19.12a
- M19.12(adj) is w/in 5% of M19.12a in 30 of 38 years, w/in 10% in 36



Interjurisdictional issues (1 of 2)

- Previous (2018, 2020) attempts at incorporating movement into the assessment model have failed to move beyond the “preliminary assessment” stage, as the Team and SSC have been skeptical of the possibility of estimating movement rates given present data limitations
- Likewise, none of the CIE reviewers recommended development of an assessment model incorporating movement
- The reviewers did make other related recommendations, however
- Unlike most previous discussions by the Team or SSC, which focused primarily on movement between the EBS and NBS, the CIE reviewers’ interest focused on movement between U.S. and Russian jurisdictions
- Development of a “simulation study” and “analytical models,” rather than an assessment model, were recommended as ways to increase understanding of the interjurisdictional issues involved, including the possibility of disproportionate harvesting

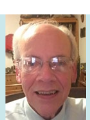
NOAA FISHERIES



Interjurisdictional issues (2 of 2)

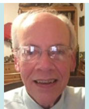
- Attachment 2.1.2 develops a very simple, deterministic, age-structured, two-area model, with results focused primarily on age-aggregated (but area-specific) equilibrium outcomes
- The primary goals are to understand which variables determine both relative and absolute biomasses and yields in the two areas, how various outcomes may be independent of specific parameters, and how various parameters covary in order to result in particular outcomes
- In general, the results illustrate the intuitive principle that, the more the stock is concentrated in the EBS/NBS, either due to recruitment being concentrated in that area, fish tending not to stray from that area once they arrive, or both—the smaller the impacts of fishing in the WBS
- Overall, reported WBS catches in recent years (Lajus et al. 2019) do not appear to be particularly high relative to estimates of WBS survey biomass (O’Leary et al. 2021)

NOAA FISHERIES



Harvest control rules and averaging (1 of 2)

- For the last three years, the senior author of the assessment and various members of the Team and SSC have spent considerable time and effort debating the issue of whether, in the context of ensemble modeling, the harvest control rules should be applied *before* or *after* model averaging
- At the request of the Team and SSC, this issue was considered yet again during the CIE review, but the responses of the reviewers were, generally speaking, somewhat nuanced
- None of the reviewers gave an unqualified endorsement of either approach, and one of them suggested that a conclusion would have to await "further investigations and examples"
- Attachment 2.1.3 was developed in response to this suggestion



Harvest control rules and averaging (2 of 2)

- A central focus of the analysis is the relative uncertainty in ABC or OFL resulting from the two procedures
- In brief, the uncertainty associated with the “before” approach is very likely to be greater, and perhaps substantially so, than the uncertainty associated with the “after” approach
- This is because the “before” approach incorporates both the *within*-model and *between*-model uncertainty in F_{ABC} or F_{OFL}, whereas the “after” approach ignores both of these
- In addition, the attachment summarizes various theoretical arguments for and against each procedure, ultimately concluding that the “before” approach is superior
- Similarly, Burnham and Anderson (2002) concluded that *parameters* in nonlinear models “should not be averaged” and that, instead, “model averaging the *expected response variable*” is the appropriate course

NOAA FISHERIES