Uncertainty in Pacific Bluefin Tuna Management Strategy Evaluation

ISC Pacific Bluefin Tuna Working Group
Present by Huihua Lee, Desiree Tommasi

Why Uncertainty?

PBF management advice has been based on stock assessments grounded in a single, best-case set of assumptions with some uncertainty considered.

Uncertainty arising from limited understanding of the true underlying system and challenges in effective implementation may dominate the sources of errors.

Management advice is based on a fully specified set of rules that will be tested through MSE simulations across diverse scenarios, explicitly accounting for uncertainty.

Punt and Donovan in 2007; Punt et al. 2014

Sources of Uncertainties

I Observations

Arising from sampling and monitoring of resources, related to collecting catch, size comp, and abundance index data

e.g., standard deviation for each observation in CPUE, effective sample size given to the size comps

II Model Assumptions

Arising from lack of knowledge of population dynamics functional forms

e.g., Beverton-Holt or Ricker SR, asymptotic or dome-shaped selectivity, different error structure of the data

III Parameter Assumptions

Arising from lack of knowledge of population dynamics parameters

e.g., natural mortality, steepness

Sources of Uncertainties

IV Process

Arising from seemingly unpredictable natural variability in population parameters affecting abundance

e.g., future recruitment, future time-varying selectivity

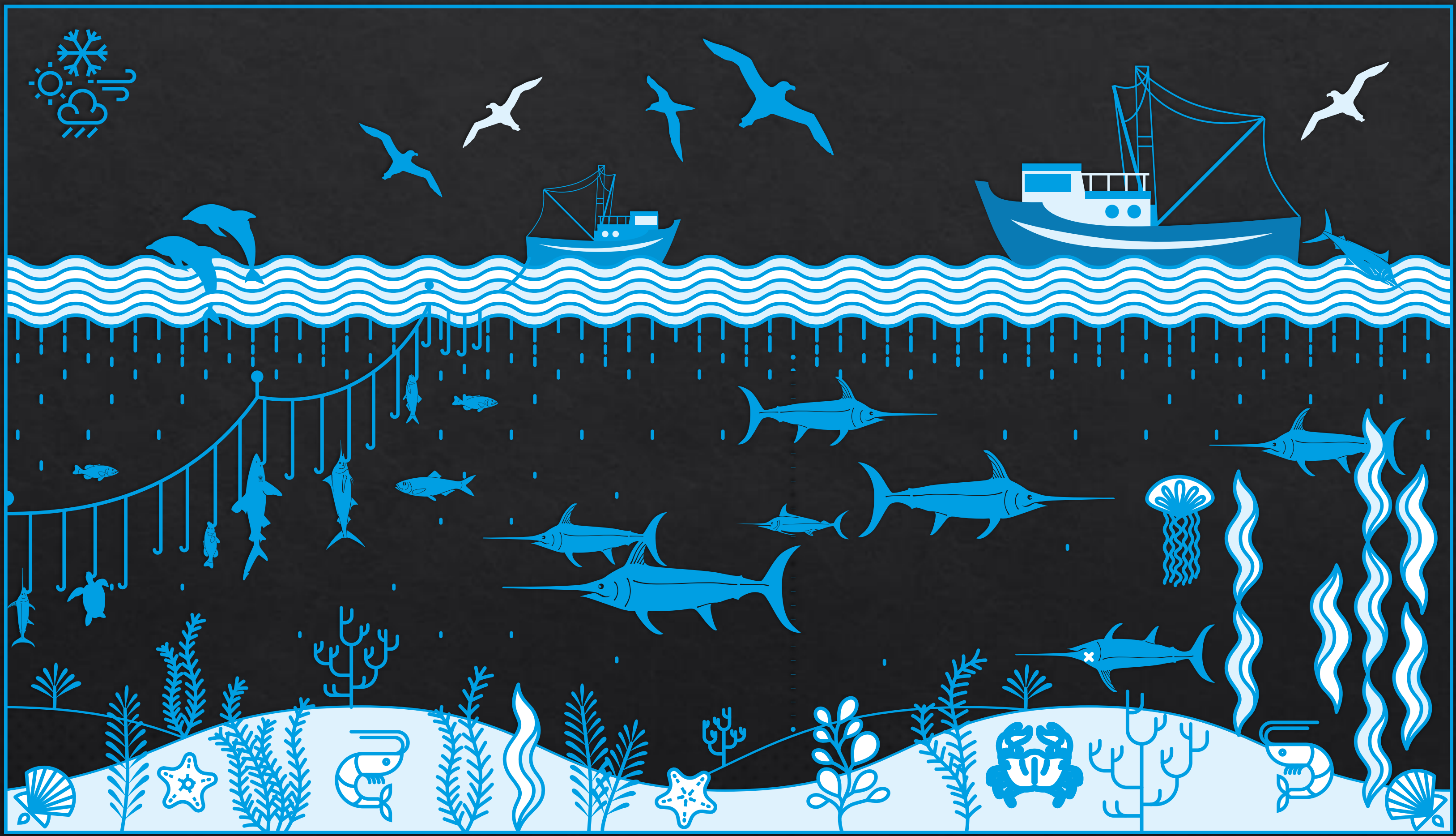
V Implementation

Arising from problems in enforcement of measures taken, imperfectly implemented management actions

e.g., catches are more than the TACs, discards exist but not included in the consideration

Example

UNCERTAINTIES PERTINENT TO INDIAN OCEAN SWORDFISH MANAGEMENT STRATEGY EVALUATION



https://edu.iotc.org/uncertainties.html

SOURCES OF UNCERTAINTY

UNCERTAINTIES CONSIDERED IN MSE ●
UNCERTAINTIES EXCLUDED FROM MSE ○

CATCH

1. Catch mis- and under-reporting
2. Discard mortality
3. Unreported discards
4. **CPUE standardisation/conflicts**
5. Bycatch
6. **Selectivity; gear selectivity/ catchability changes by fleet (e.g. gear/equipment changes)**
7. Changes in effort distribution: seasonal dynamics (stock/fleet)

LIFE HISTORY TRAITS

10. **Growth and maturity**
11. **Natural mortality (M)**
12. Sex dependent migration: spatial sexual segregation of the stock (real or observed)
13. Fecundity
14. Stock structure and mixing; group dynamics, skipped-spawning, density dependence

ENVIRONMENTAL

8. Climate change and/or increased variability's potential to change population dynamics
9. Environmental forcing; environmental considerations and behaviour

MODEL

15. Model complexity
16. **Steepness**
17. **Alternative data weights (length comp); length compositions effective sample size**
18. **Scaling**

SOCIO-ECONOMIC

19. Economic uncertainty; market and other economic data to be used in assessing the risks
20. Uncertainty over objectives; management objectives
21. Uncertainty over reference points; lack of information on virgin stock levels
22. Risk attitudes of managers
23. **Catchability increase**
24. Effect of regulations on effort; minimum size recommendation; implementation options
25. Social impacts on local communities; impacts/effect on small local communities
26. Illegal fishing; regulations that change the balance of effort between legal and illegal fisheries
27. Effect of regulations on species; impacts and effect on global distribution of the species.

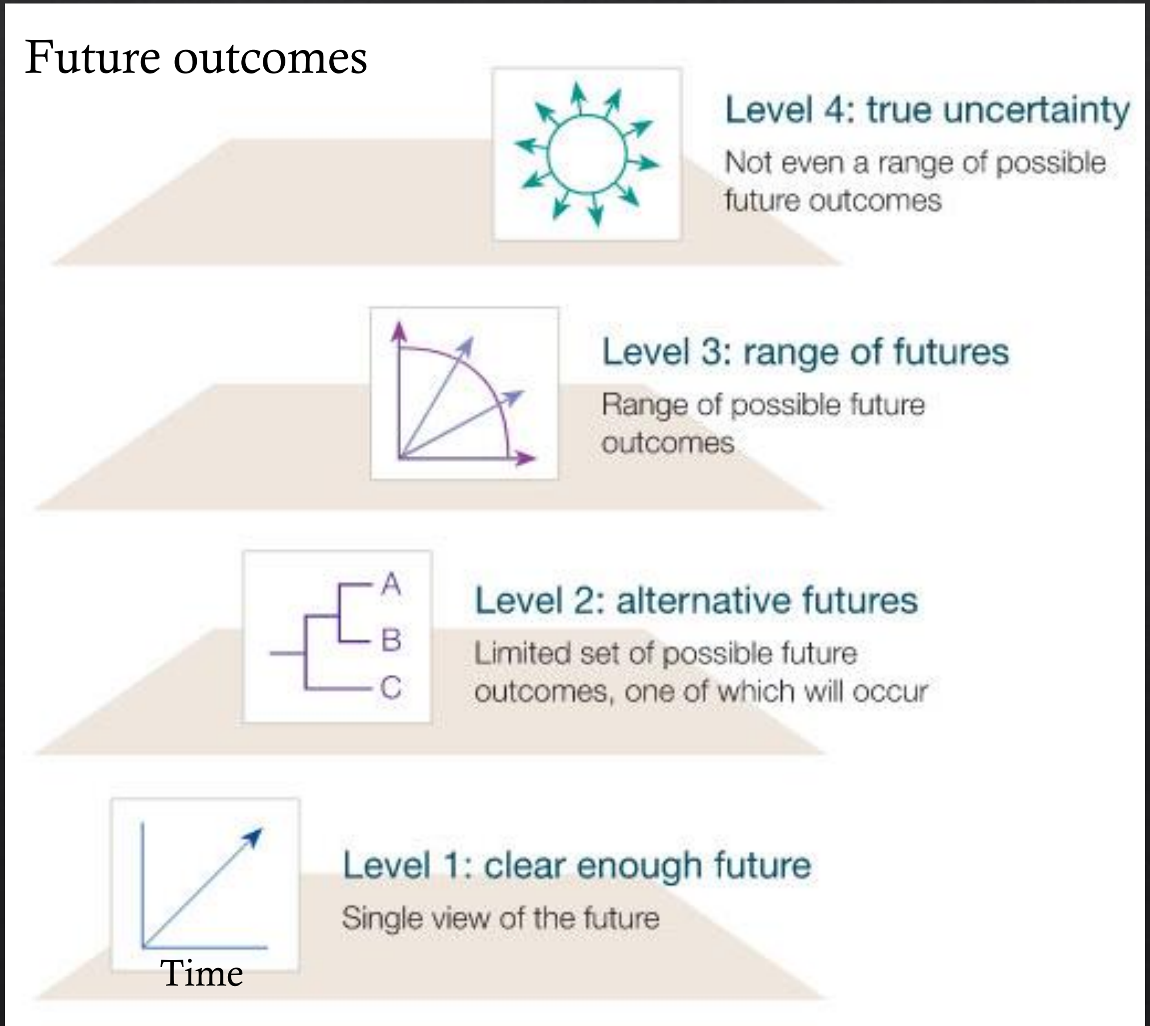
POPULATION STRUCTURE

28. Oxygen minimum zone, i.e. vertical displacement of individuals
29. Cyclic movement of adult swordfish
30. Changes in migration; environmental factors that influence migration patterns
31. Spatio-temporal dynamics of sub-populations
32. Existence of genetically distinct and vulnerable sub-stocks
33. Sex ratio
34. Interactions with other species
35. **Recruitment Variability Recruitment failure of success (cyclic trends/regime shift)**

REFERENCE POINTS

36. Dynamics of reference points; stationarity, cohort year effects, density dependence

Making The Most of Uncertainty



Courtney (2001), McKinsey & Company

- While it is clearly desirable to conduct trials for all combinations for the levels for each factor, this is often computationally impossible.

Making The Most of Uncertainty: Best Practices

Table 3 List of factors, whose uncertainty commonly has a large impact on management strategy performance, which should be considered for inclusion in any management strategy evaluation.

Productivity

- Form and parameters of the stock–recruitment relationship.
- Presence of depensation.
- Extent of variation and correlations in recruitment about the stock–recruitment relationship.
- Occasional catastrophic mortality or recruitment events.

Non-stationarity

- Changes in the stock–recruitment relationship.
- Time-varying natural mortality (potentially a multispecies operating model).
- Time-varying carrying capacity (regime shift; linked to environmental variables or multispecies effects).
- Time-varying growth and selectivity.

Other factors

- Spatial and stock structure.
- Technical interactions.
- Time-varying selectivity, movement and growth.
- Initial stock size (unless it is estimated reliably when conditioning the operating model).

Data-related issues

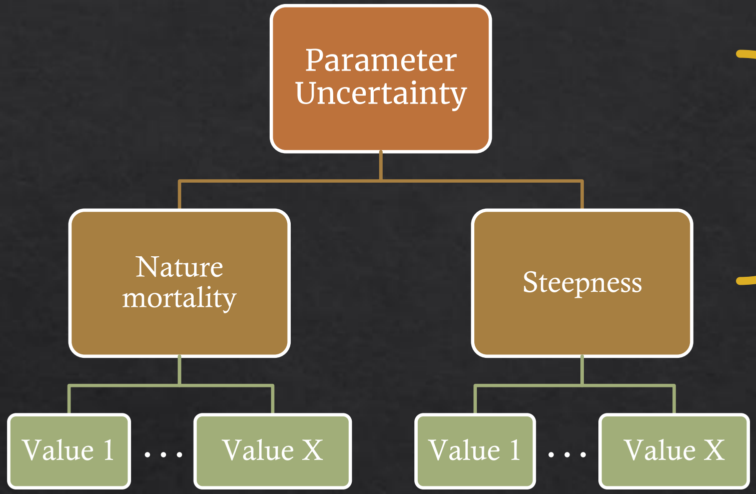
- CVs and effective samples sizes of data.
- Changes in the relationship between catchability and abundance.
- Changes in survey bias (fishery-independent data).
- Survey and sampling frequency.
- Ageing error.
- Historical catch inaccuracy (bias).

Outcome (Implementation) uncertainty

- Decision-makers adjust or ignore management advice.
- Realized catches differ from total allowable catches due to mis-reporting, black market catches, discards, etc.

Punt et al. 2014

Making The Most of Uncertainty for PBF



1. Identify the set of factors

   Factors are perceived to contribute to the uncertainty for PBF

2. Select levels for each factor

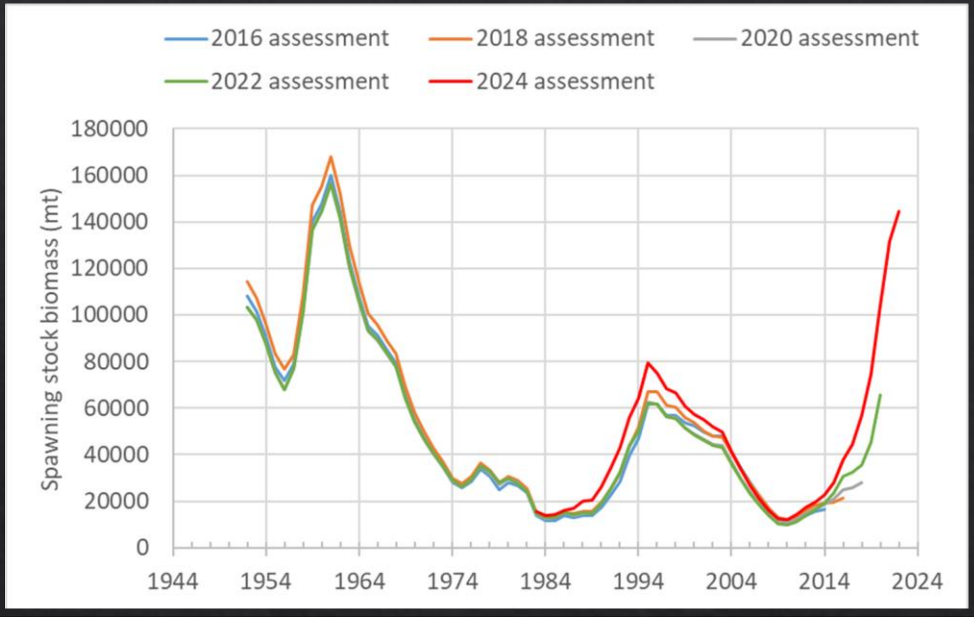
3. Select plausible combinations of factors

4. Triage uncertainty

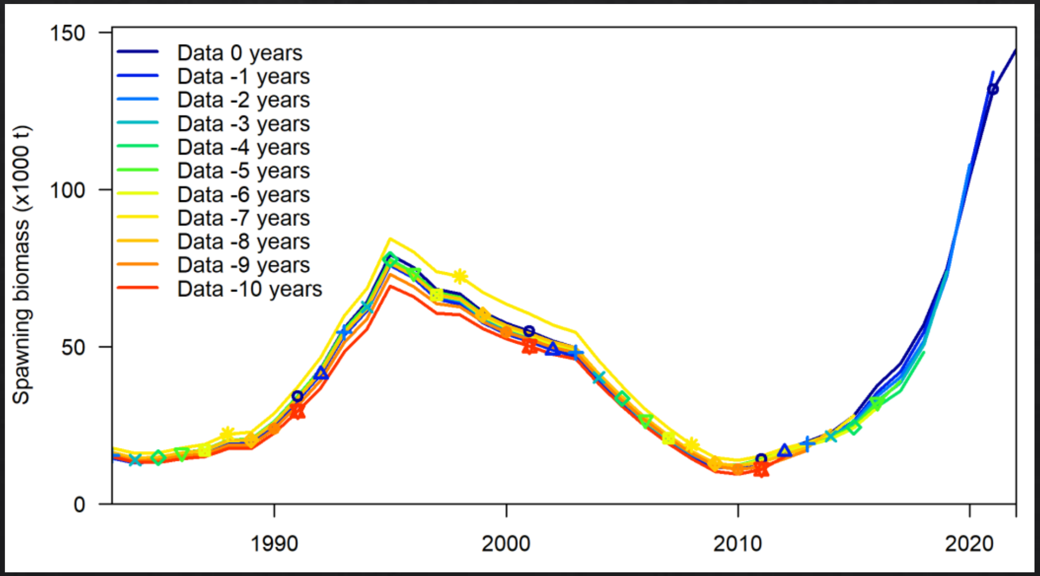
   Divide MSE trials into a "reference set" and a "robustness set".

   A reference set reflects the most plausible hypotheses.
   A robustness set reflects fairly unlikely hypotheses, even though they are still plausible.

1 Identify Factors: Observation Uncertainty



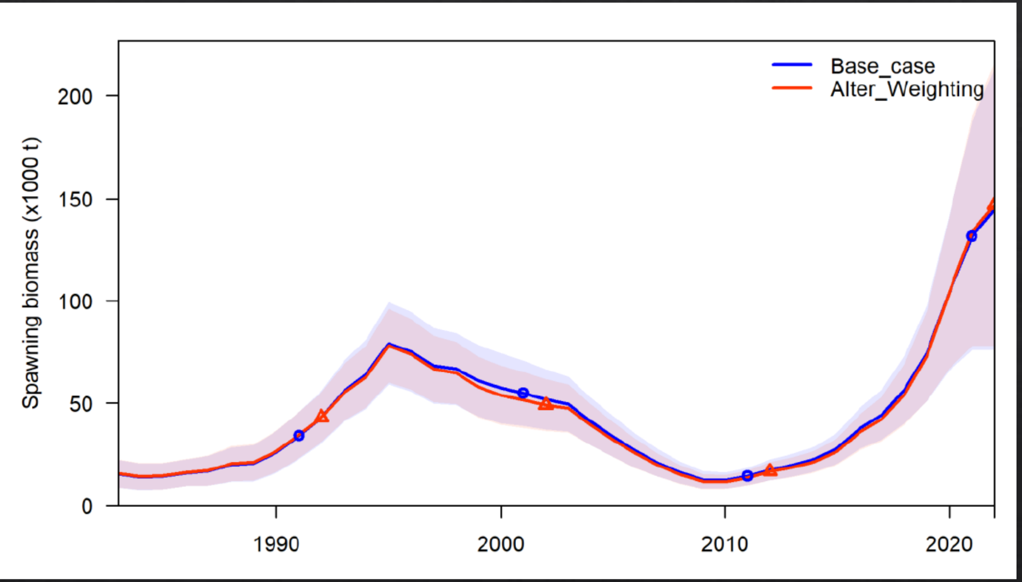
- Major model assumptions are similar.
- Consistent results were found among stock assessments.



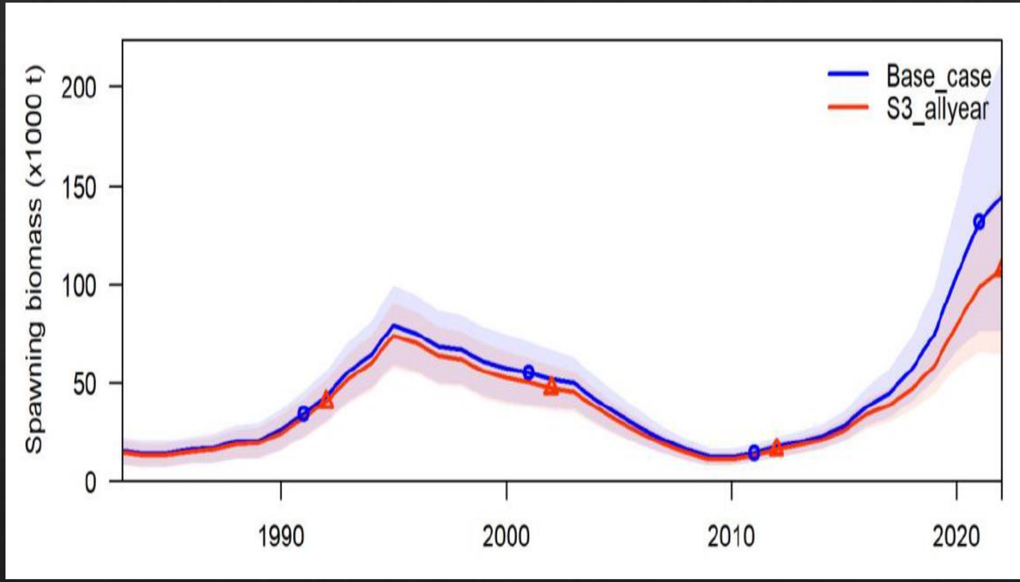
- Little retrospective pattern were found.
- When analyzing data over time, there is no consistent trend where the results significantly change; essentially, the data appears stable and doesn't show a pattern of altering conclusions based on the data points considered.

1 Identify Factors: Observation Uncertainty

Alternative data weighting

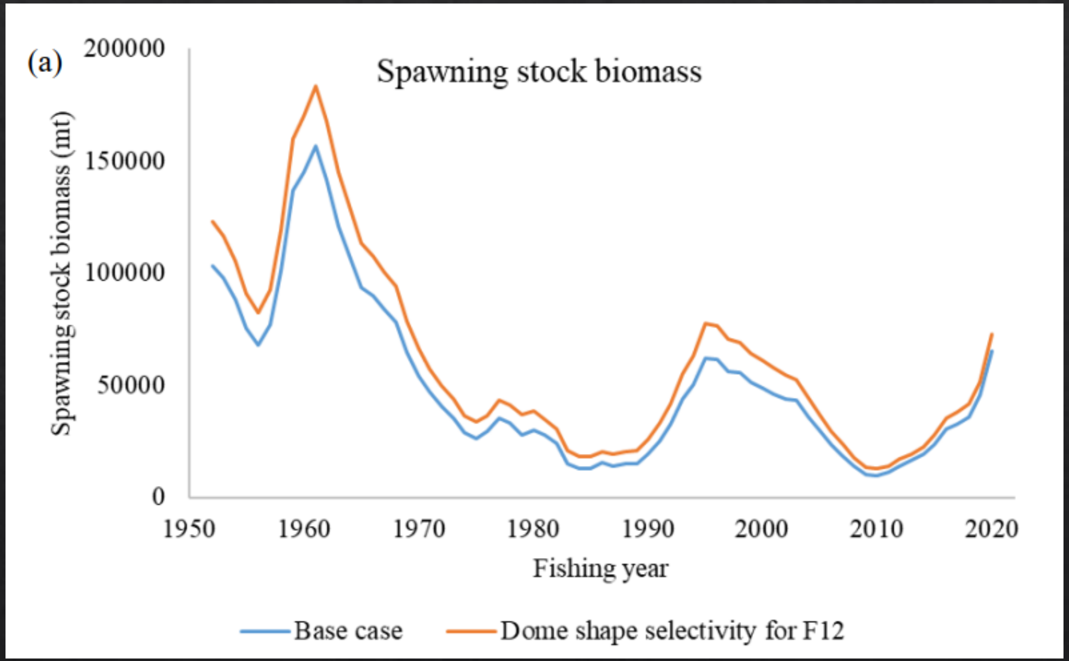


Troll index for 1983-2016



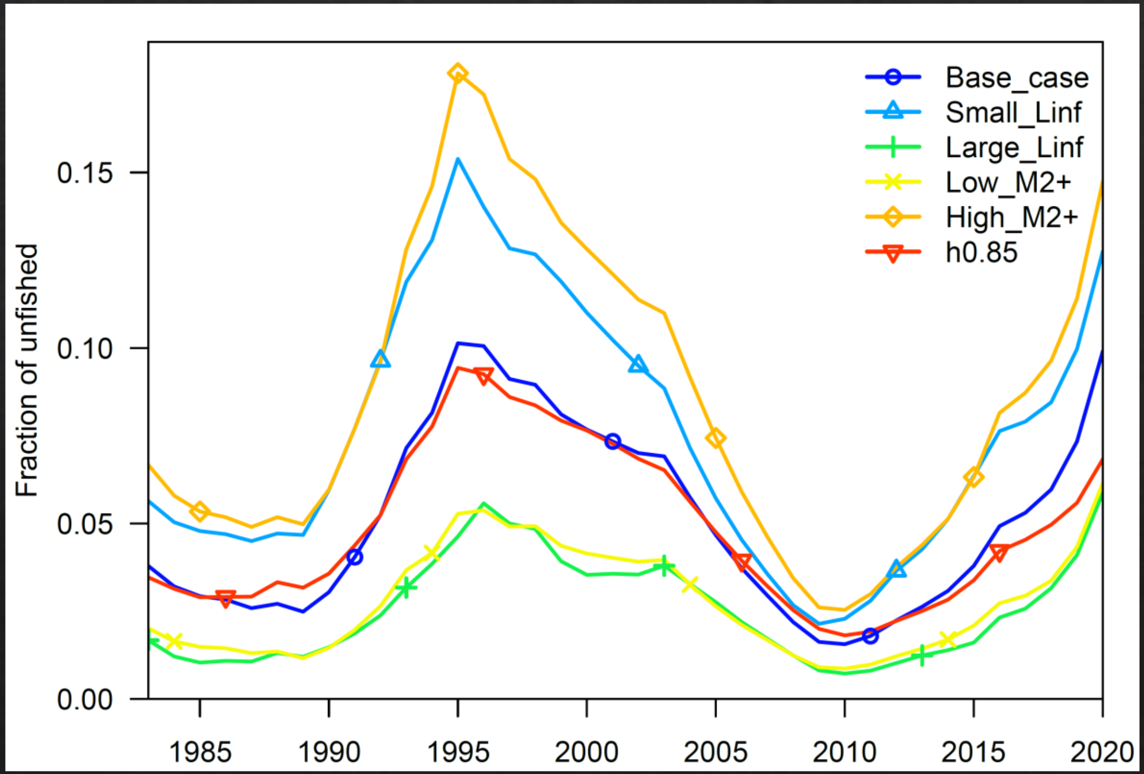
1 Identify Factors: Model Uncertainty

Alternative selectivity forms for Taiwan longline fishery

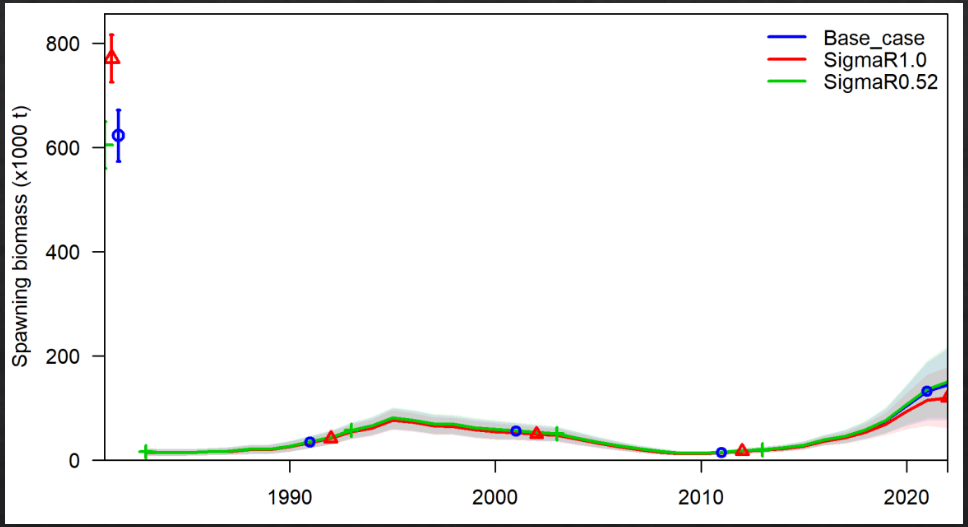


1 Identify Factors: Parameter Uncertainty

Productivity parameters: Natural mortality, Growth, steepness



Recruitment variability



1 Identify Factors: Summary

- Parameter uncertainty related to productivity is the most influential among Observation, Model, and Parameter uncertainties and should encompass the wildest plausible range for PBF.
- MSE simulations will handle process and implementation uncertainties (more details below).

2 Select Levels for Each Factor: Natural Mortality

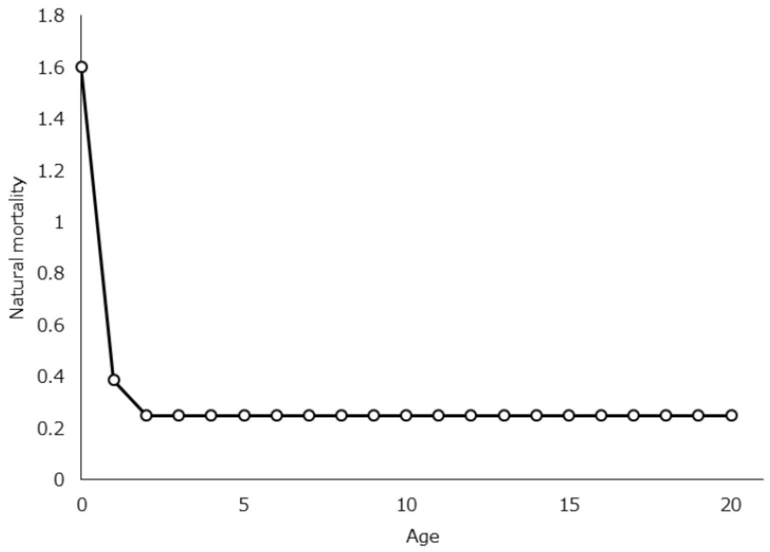


Figure 2-5. Assumed natural mortality (*M*) at age of Pacific bluefin tuna (*Thunnus orientalis*) used in this stock assessment.

- Age 0 fish was from a conventional tagging study.
- Age-1 fish was based on length-adjusted M estimated from conventional tagging studies on southern bluefin tuna.
- Age-2 and older fish was from the median value obtained across a suite of empirical and life-history based methods.

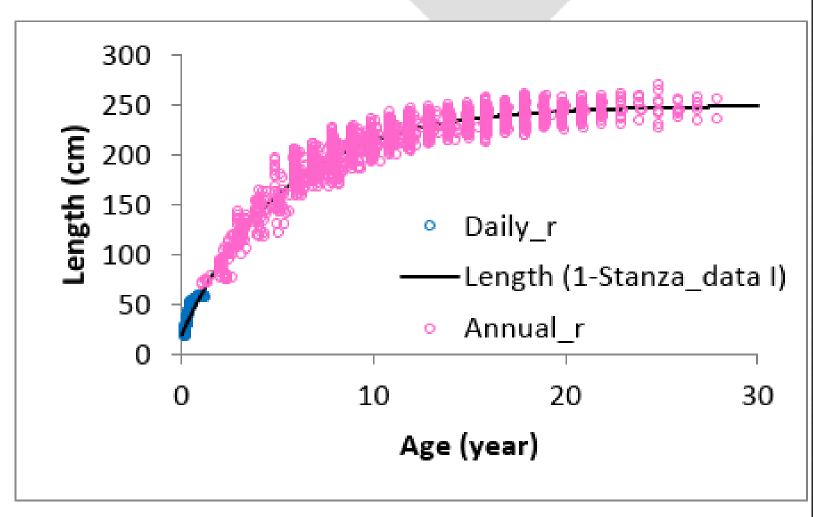
2 Select Levels for Each Factor: Adult Natural Mortality

- Maunder et al. (2023) reviewed methods for examining *M* and recommended focusing on the maximum observed age (*tmax*) as it provides a more direct relationship with *M*.

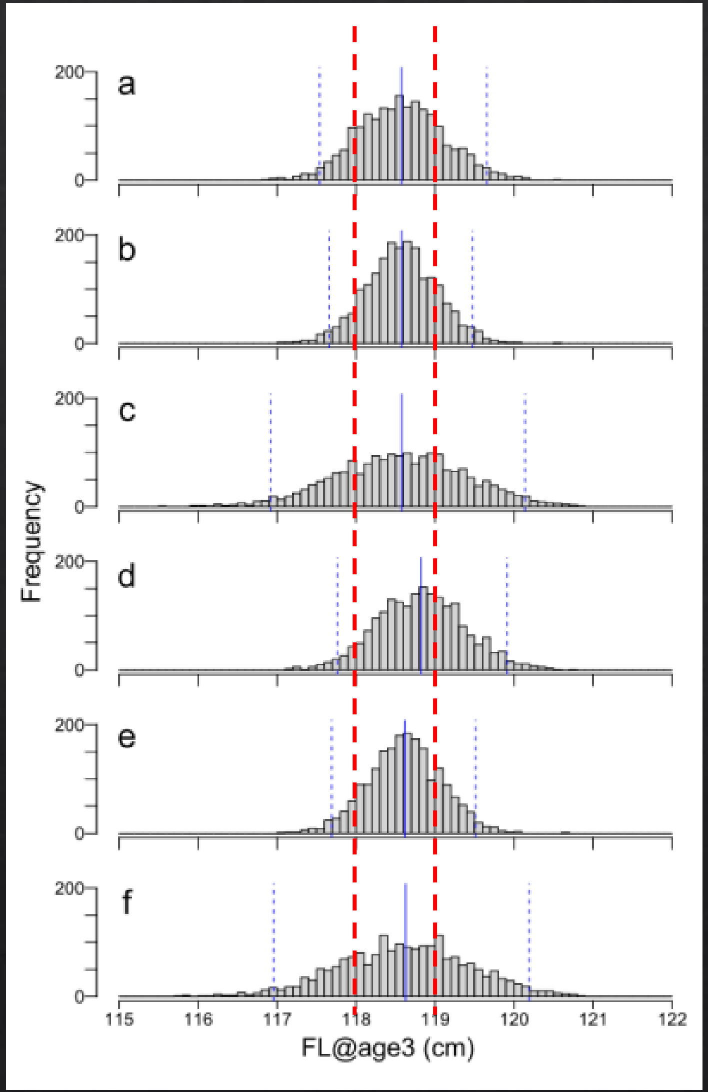
Method	M1	M2	Parameters input	Reference
Then_nls	**0.256**	**0.231**	**maximum age**	**Then et al. (2015)**
Then_lm	**0.216**	**0.192**	**maximum age**	**Then et al. (2015)**
Hamel_Amax	**0.216**	**0.193**	**maximum age**	**Hamel (2015)**
ZM_CA_pel	0.162	0.131	maximum age, k, t0	Alverson and Carney (1975) Zhang and Megrey (2006)
Then_VBGF	0.196		Linf, k	Then et al. (2015)
Hamel_k	0.33		k	Hamel (2015)
Jensen_k1	0.282		k	Jensen (1996, 1997)
Jensen_k2	0.301		k	Jensen (1996, 1997)
Roff	0.362		k, age at maturity	Roff (1984)
Jensen_Amat	0.33		age at maturity	Jensen (1996, 1997)
Ri_Ef_Amat	0.317		age at maturity	Rikhter and Efanov (1976)

- We explored a range of M_2^+ values, 0.193 and 0.25 year^{-1} corresponding to 22 and 28 years *tmax*.

2 Select Levels for Each Factor: Growth



Bootstrapping



- Annuli data for 1,782 fish

(70.5-271 cm in FL, 1-28 yrs old)

- Daily increment data for 228 fish

(18.6-60.1 cm in FL, 51-453 days old after hatching)

Fukuda et al. 2015
ISC/15/PBFWG-2/11

Ishihara et al. 2023

- The median of estimated length-at-age 3 ranged from 118.57 to 118.82.
- The 95% confidence interval for the estimated length-at-age 3 was within ±2 cm from the median.
- We consequently selected the range of length-at-age 3 spanning from 118 to 119 cm corresponding to 248.6 to 250.9 cm L_{inf}.

3 Select Plausible Combination of productivity parameters

$$M_{2+} * h * L_2$$

M at age 2 +   Length at age 3

- We reviewed the direct data and indirect information to initially select range of the parameters
- Next, we fit fishery data into the assessment model using the alternative productivity assumptions and conducted diagnostic analyses for each grid.

3 Select Plausible Combination of productivity parameters

		M_{2+}=0.193			M_{2+}=0.25		
		L_2=118 (L_{inf}=248.6)	L_2=118.57 (L_{inf}=249.9)	L_2=119 (L_{inf}=250.9)	L_2=118 (L_{inf}=248.6)	L_2=118.57 (L_{inf}=249.9)	L_2=119 (L_{inf}=250.9)
Steepness	0.81						
	0.83						
	0.85						
	0.87						
	0.89						
	0.91						
	0.93						
	0.95						
	0.97						
	0.99						
	0.999						

Grid / Grid model
Alternative grid

Base grid

Lee and Tommasi 2024
ISC/24/PBFWG-1/15

Full dynamics model with various productivity assumptions and diagnostics

1. *Convergence*

   Jitter and positive-definite Hessian matrix

2. *Goodness-of-fit*

   Negative log-likelihood (NLLs)

3. *R_0 likelihood profile*

   conflicts arose among these sources

4. *Retrospective analyses*

   identify patterns that occurred in the historical data and model structure

5. *Age-structured production model with recruitment (ASPM-R)*

   evaluate the current state of the production function

1. Convergence

percentage of runs resulting in a positive-definite Hessian matrix in jitter analyses

⟐ Purpose: Evaluate convergence towards a global minimum

		M_{2+}=0.193			**M_{2+}=0.25**		
		L_2=118 (L_{inf}=248.6)	L_2=118.57 (L_{inf}=249.9)	L_2=119 (L_{inf}=250.9)	L_2=118 (L_{inf}=248.6)	L_2=118.57 (L_{inf}=249.9)	L_2=119 (L_{inf}=250.9)
h	0.91	N/A	N/A	N/A	62%	58%	50%
	0.93	N/A	N/A	N/A	73%	81%	65%
	0.95	0%	0%	0%	88%	77%	65%
	0.97	8%	4%	0%	77%	92%	81%
	0.99	85%	62%	58%	92%	88%	92%
	0.999	92%	92%	81%	92%	**96%**	92%

Any grid with 0% of runs resulting in a positive-definite Hessian matrix will not be considered in the subsequent diagnostics and will be given a score of zero.

Lee and Tommasi 2024
ISC/24/PBFWG-1/15

2. *Goodness-of-fit*

Negative log-likelihood (NLLs)

◈ **Purpose: Evaluate goodness-of-fit given the data and model structure**

Indices

	M_{2+}=0.193			M_{2+}=0.25		
	L_2=118 (L_{inf}=248.6	L_2=118.57 (L_{inf}=249.9	L_2=119 (L_{inf}=250.9	L_2=118 (L_{inf}=248.6	L_2=118.57 (L_{inf}=249.9	L_2=119 (L_{inf}=250.9
0.91	N/A	N/A	N/A	-83	-83	-83
0.93	N/A	N/A	N/A	-82	-83	-84
0.95	.	.	.	-83	-84	-84
0.97	-83	-83	.	-83	-84	-85
0.99	-82	-84	-83	-83	-84	-85
0.999	-83	-83	-84	-83	**-85**	-85

Index data components suggest that most grids performed similarly to the base grid.

Size compositions

	M_{2+}=0.193			M_{2+}=0.25		
	L_2=118 (L_{inf}=248.6	L_2=118.57 (L_{inf}=249.9	L_2=119 (L_{inf}=250.9	L_2=118 (L_{inf}=248.6	L_2=118.57 (L_{inf}=249.9	L_2=119 (L_{inf}=250.9
0.91	N/A	N/A	N/A	1321	1310	1316
0.93	N/A	N/A	N/A	1305	1310	1314
0.95	.	.	.	1305	1309	1315
0.97	1319	1331	.	1305	1309	1316
0.99	1304	1309	1313	1305	1309	1314
0.999	1304	1308	1313	1305	**1309**	1314

The size compositions indicate that, as L_2 decreased, more grids with a fit similar to or better than the base grid were achieved.

The index and size composition components provided inconsistent grid profiles; therefore, goodness-of-fit will not be considered in the selection process.

Lee and Tommasi 2024
ISC/24/PBFWG-1/15

3. R_0 likelihood profile

◈ **Purpose: Evaluate which data sources provided information on a global scale and where conflicts arose among these source**

		M_{2+}=0.193			M_{2+}=0.25		
		L_2=118 (L_{inf}=248.6)	L_2=118.57 (L_{inf}=249.9)	L_2=119 (L_{inf}=250.9)	L_2=118 (L_{inf}=248.6)	L_2=118.57 (L_{inf}=249.9)	L_2=119 (L_{inf}=250.9)
h	0.91	N/A	N/A	N/A	Indices & Size	Indices & Size	Indices & Size
	0.93	N/A	N/A	N/A	Size	Size	Size
	0.95	.	.	.	Size	Size	Size
	0.97	Size	Size	.	Size	None	Size
	0.99	Size	Size	Size	Size	None	Size
	0.999	Size	Size	Size	Size	**Size**	Size

For the base grid, size components provided consistent estimates of the global scale (ln(R_0)).

Any grid lacking the same consistency as in the base grid will be given a score of zero.

Lee and Tommasi 2024
ISC/24/PBFWG-1/15

4. *Retrospective analyses*

Mohn's ρ

- **Purpose: Identify patterns that occurred in the historical data and model structure**

Mohn's ρ	M_{2+}=0.193			M_{2+}=0.25		
	L_2=118 (L_{inf}=248.6)	L_2=118.57 (L_{inf}=249.9)	L_2=119 (L_{inf}=250.9)	L_2=118 (L_{inf}=248.6)	L_2=118.57 (L_{inf}=249.9)	L_2=119 (L_{inf}=250.9)
0.91	N/A	N/A	N/A	-0.09	-0.11	-0.11
0.93	N/A	N/A	N/A	-0.09	-0.09	-0.08
0.95	.	.	.	-0.07	-0.06	-0.05
0.97	-0.04	-0.02	.	-0.04	-0.02	-0.02
0.99	0.05	0.05	0.08	-0.01	0	0.01
0.999	0.07	0.1	0.12	0	**0.01**	0.03

Any grid with an absolute Mohn's ρ value larger than 0.1 will be given a score of zero.

Lee and Tommasi 2024
ISC/24/PBFWG-1/15

5. Age-structured production model with recruitment (ASPM-R)

evaluate the current state of the production function

⟐ Purpose: Evaluate the current state of the production function

NLLs		M_{2+}=0.193			M_{2+}=0.25		
h		L_2=118 (L_{inf}=248.6)	L_2=118.57 (L_{inf}=249.9)	L_2=119 (L_{inf}=250.9)	L_2=118 (L_{inf}=248.6)	L_2=118.57 (L_{inf}=249.9)	L_2=119 (L_{inf}=250.9)
	0.91	N/A	N/A	N/A	-39.9	-39.1	-40.2
	0.93	N/A	N/A	N/A	-39.8	-40.6	-41.2
	0.95	.	.	.	-40.3	-41.3	-42.0
	0.97	-39.3	-39.6	.	-41.4	-42.2	-43.2
	0.99	-38.8	-39.8	-39.1	-43.5	-44.7	-45.3
	0.999	-40.5	-40.8	-41.1	-43.8	**-44.7**	-45.4

The NLLs generally deteriorated when *h* was smaller than the base value, regardless of M_2^+ or L_2 values. The selected range of *h* expanded when either M_2^+ or L_2 was larger.

Any grid displaying a statistically significant degradation in NLLs, thus hindering the production relationship, will be given a score of zero.

Lee and Tommasi 2024
ISC/24/PBFWG-1/15

3 Select Plausible Combination of productivity parameters

Ensemble diagnostics scores

(jitter, R_0 profile, retrospective, and ASPM-R analyses)

		M_{2+}=0.193			M_{2+}=0.25		
		L_2=118 (L_{inf}=248.6)	L_2=118.57 (L_{inf}=249.9)	L_2=119 (L_{inf}=250.9)	L_2=118 (L_{inf}=248.6)	L_2=118.57 (L_{inf}=249.9)	L_2=119 (L_{inf}=250.9)
h	0.91	N/A	N/A	N/A	3	2	2
	0.93	N/A	N/A	N/A	3	3	3
	0.95	0	0	0	3	3	3
	0.97	3	3	0	3	2	4
	0.99	3	3	3	4	3	4
	0.999	3	2	2	4	4	4

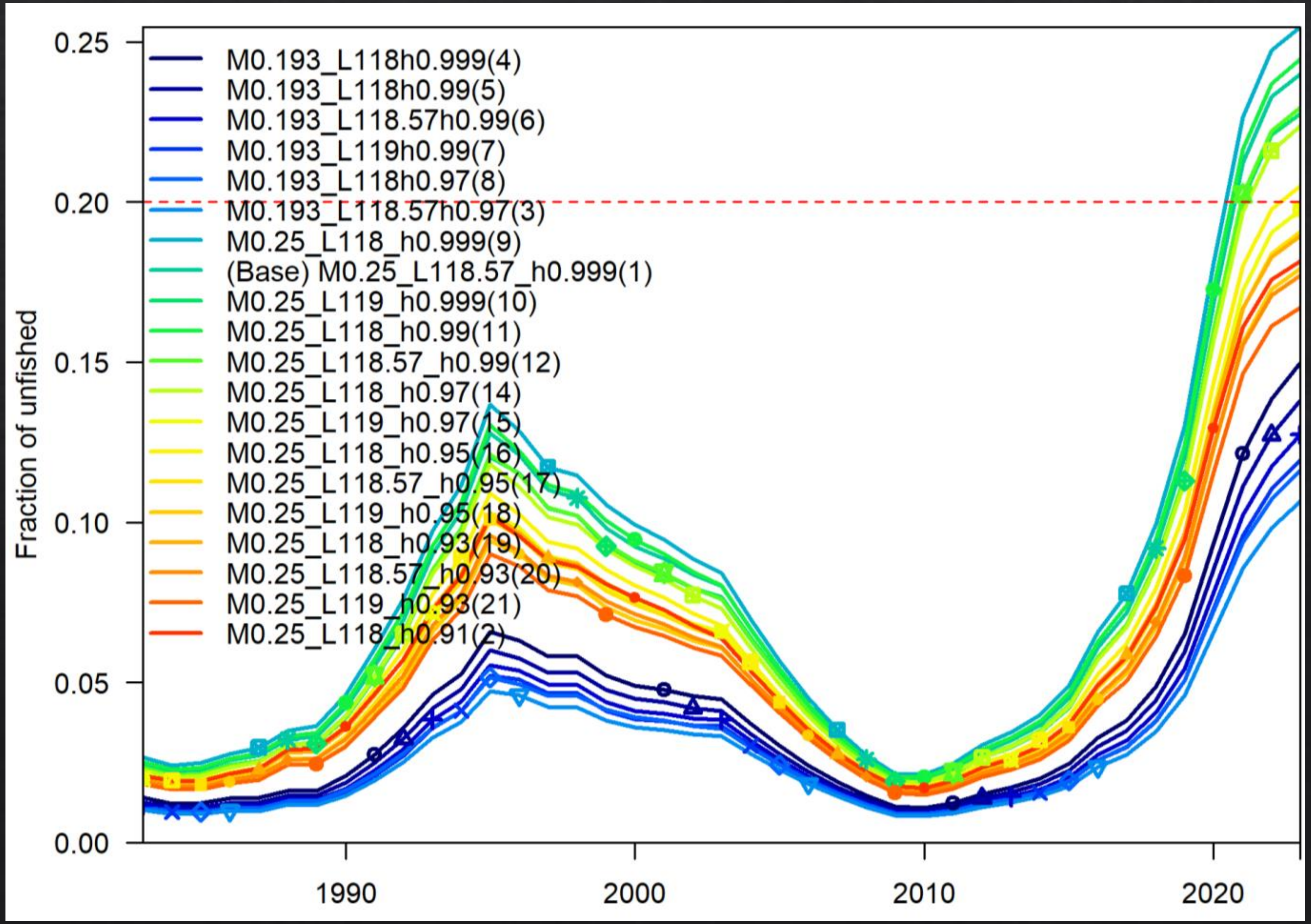
The scores range from 0 to 4. The scores reveal conflicting information across retrospective analyses, R_0 profile, and ASPM-R. Specifically, ASPM-R favored higher values for M_2^+ and *h*, while R_0 profile leaned towards lower values for *h*.

Grids that passed three or more diagnostics were recommended.

Lee and Tommasi 2024 ISC/24/PBFWG-1/15

Reference Set

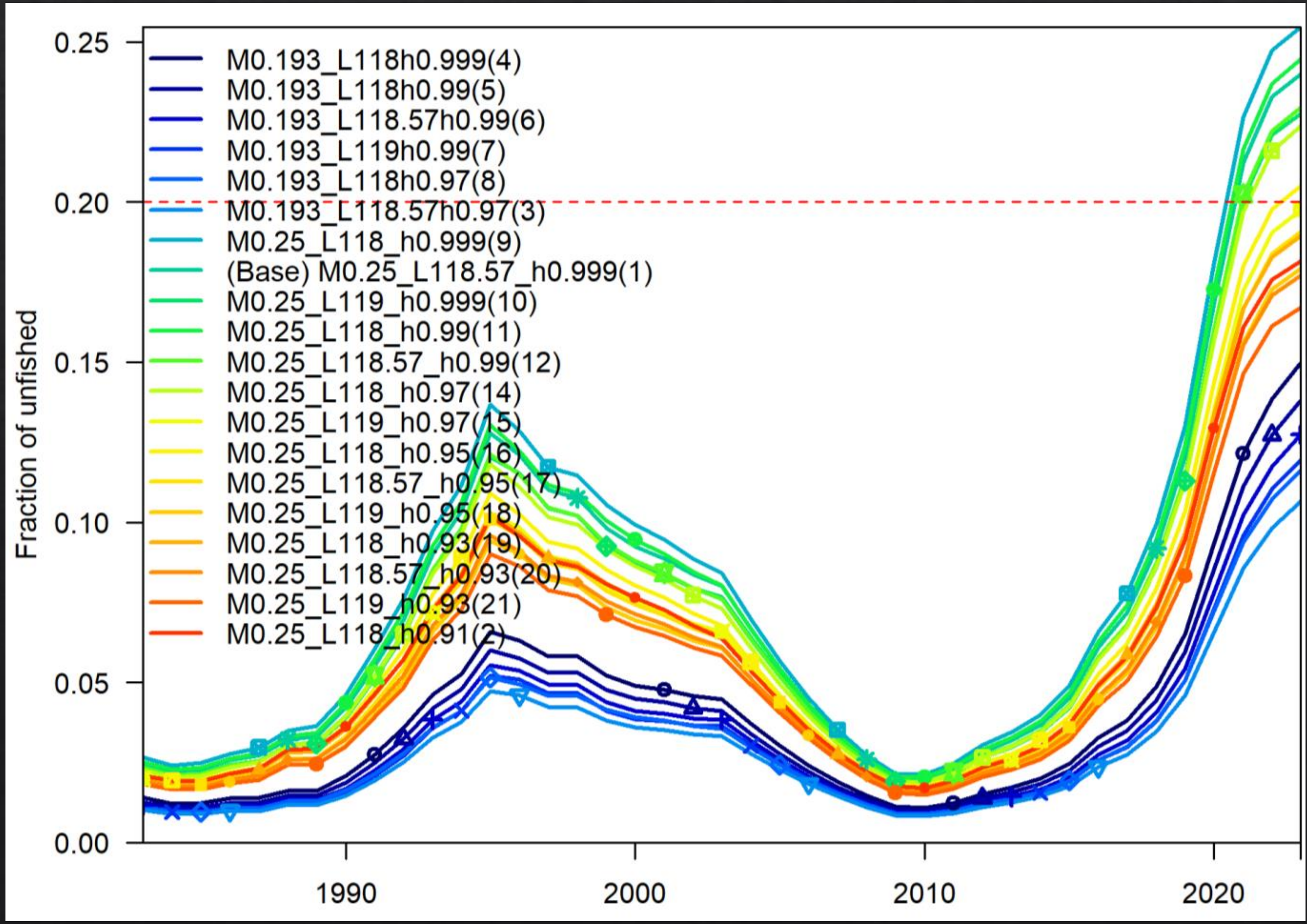
- A set of operating models is a representation of the “truth” and can attempt to account for major sources of uncertainties in the system.



OM#	M^{2+}	L at age 3	h
1	0.25	118.57	0.999
2	0.25	118	0.91
3	0.193	118.57	0.97
4	0.193	118	0.999
5	0.193	118	0.99
6	0.193	118.57	0.99
7	0.193	119	0.99
8	0.193	118	0.97
9	0.25	118	0.999
10	0.25	119	0.999
11	0.25	118	0.99
12	0.25	118.57	0.99
13	0.25	119	0.99
14	0.25	118	0.97
15	0.25	119	0.97
16	0.25	118	0.95
17	0.25	118.57	0.95
18	0.25	119	0.95
19	0.25	118	0.93
20	0.25	118.57	0.93
21	0.25	119	0.93

Reference Set

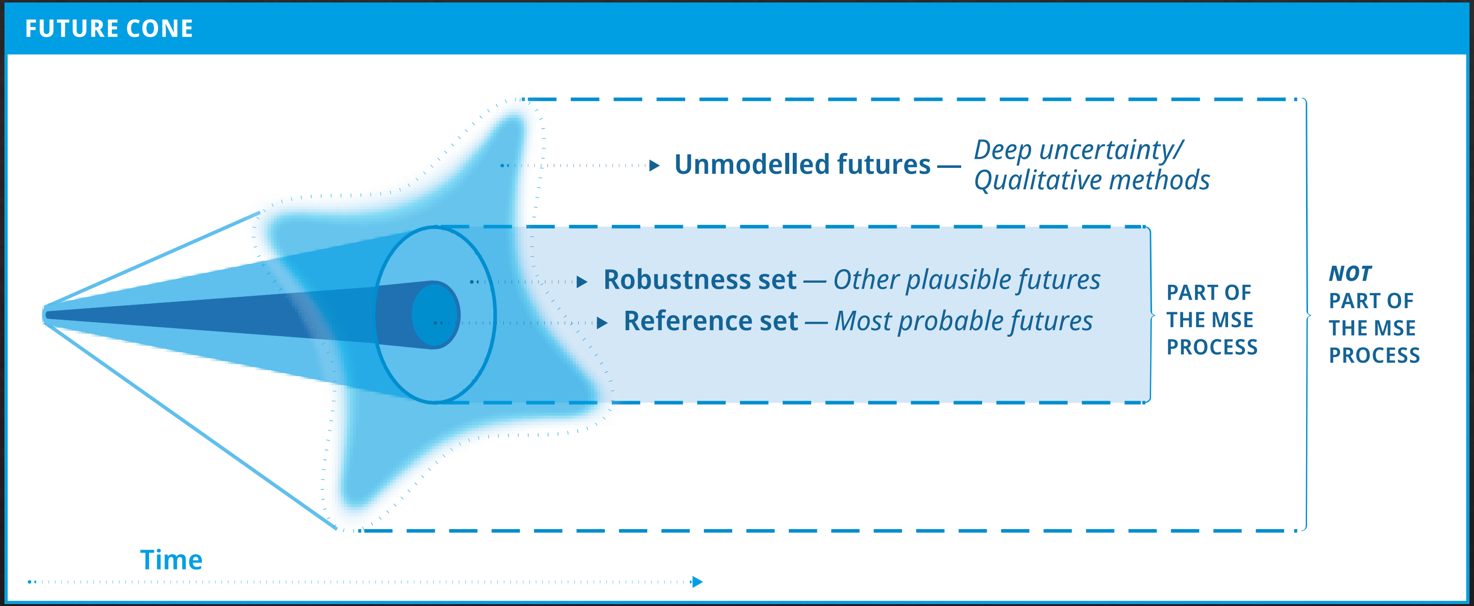
- In initial trial MSE runs with perfect estimation, OM8 was not stable and was removed. This model also had few converged runs in the jitter diagnostic.



OM#	M^{2+}	L at age 3	h
1	0.25	118.57	0.999
2	0.25	118	0.91
3	0.193	118.57	0.97
4	0.193	118	0.999
5	0.193	118	0.99
6	0.193	118.57	0.99
7	0.193	119	0.99
~~8~~	~~0.193~~	~~118~~	~~0.97~~
9	0.25	118	0.999
10	0.25	119	0.999
11	0.25	118	0.99
12	0.25	118.57	0.99
13	0.25	119	0.99
14	0.25	118	0.97
15	0.25	119	0.97
16	0.25	118	0.95
17	0.25	118.57	0.95
18	0.25	119	0.95
19	0.25	118	0.93
20	0.25	118.57	0.93
21	0.25	119	0.93

4 Reference Set vs Robustness Set

Future Cone / Imagining Possible Futures: reference scenarios, robustness scenarios, and deep uncertainty.



https://edu.iotc.org/operating-models.html

- The reference set for PBF is generated by 20 OMs that passed the plausibility test. These were resampled to generate 100 simulated worlds.
- Three robustness set for PBF
  - Doubling of discards
  - Catchability change in TWLL (2% annual increase)
  - Effect of climate change – 10 year recruitment drop like in the 1980s