RISK ANALYSIS FOR BIGEYE TUNA: model weights

Mark N. Maunder, Haikun Xu, Carolina Minte-Vera, Cleridy E. Lennert-Cody, Juan L. Valero, Alexandre Aires-da-Silva

11TH MEETING SCIENTIFIC ADVISORY COMMITTEE San Diego, California (USA)
11-15 May 2020
Postponed until a later date to be determined

The staff's pragmatic risk analysis approach

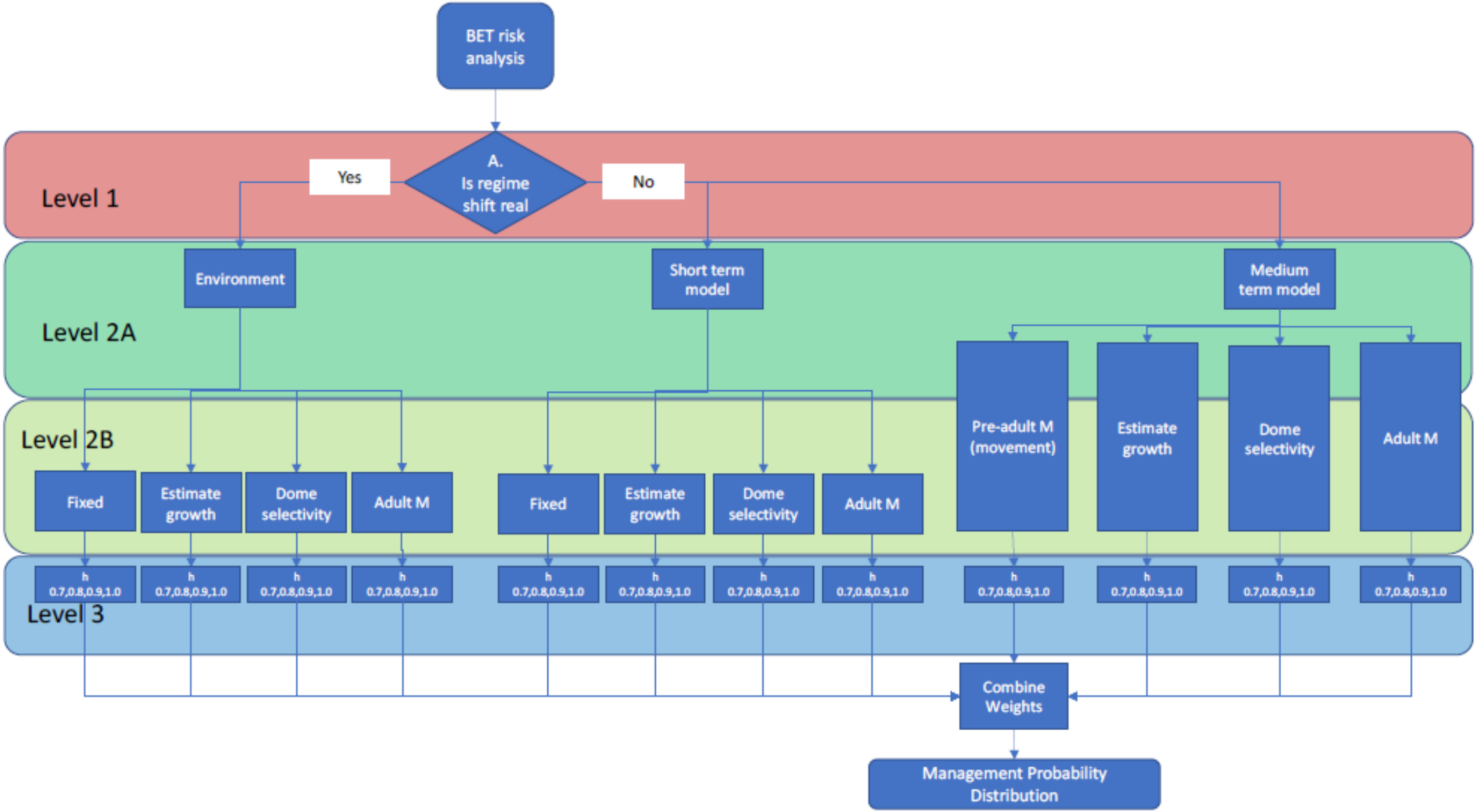
Described in Maunder et al. 2020 (SAC-11- INF-F):

1. **Identify alternative hypotheses** ('*states of nature*') about the population dynamics of the stock that address the main issues in the assessments
   - YFT: SAC-11-J; BET: SAC-11 INF-F
2. **Implement stock assessment models representing alternative hypotheses**
   - YFT: SAC-11-07; BET: SAC-11-06
3. **Assign relative weights to each hypothesis (model)**
   - YFT: SAC-11 INF-J; BET: SAC-11 INF-F
4. **Compute combined probability distributions for management quantities using model relative weights**
   - SAC-11-08

List of models *retained* in the risk analysis

Model name	Number	Description	Note
Env-Fix	1	Environment, Fixed	
Env-Gro	2	Environment, Estimate growth	
Env-Sel	3	Environment, Dome selectivity	
Env-Mrt	4	Environment, Adult mortality	
~~**Rcr**~~		~~Ricker~~	**Not shown (model does not converge)**
~~**Ind**~~		~~Index not representative~~	**Not shown (model weight=0)**
Srt-Fix	5	Short-term, Fixed	
Srt-Gro	6	Short-term, Estimate growth	
Srt-Sel	7	Short-term, Dome selectivity	
Srt-Mrt	8	Short-term, Adult mortality	
Mov	9	Pre-adult movement	
Gro	11	Estimate growth	
Sel	11	Dome selectivity	
Mrt	12	Adult mortality	
~~**Cmp**~~		~~Unrepresentative longline composition~~	**Not shown (model weight=0)**

Flow chart for bigeye tuna



Assigning model weights

- Model are weighted by the IATTC staff (stock assessment authors)

- Weight categories
    - None: 0
    - Low: 0.25
    - Medium: 0.5
    - High: 1
- Scaled weights: sum to 1

Assigning model weights

- Level 1 (regime shift hypothesis) is weighted independently based solely on expert opinion
- Level 2 is weighted based on several criteria:
  - Expert opinion
  - Convergence
  - Fit to data
  - Plausible parameter estimates
  - Plausible model results
  - Model diagnostics
  - Recruitment shift metric
  - Empirical selectivity vs. estimated selectivity
- Level 3 (steepness hypothesis) is weighted independently based solely on expert opinion

W(regime shift) – level 1

Consensus among experts: the weights of the recruitment regime shift being real and not real are low and high, respectively

“There have been some physical and biological changes in the pelagic EPO, but their timing and magnitude does not necessarily correspond to the increase in bigeye recruitment and similar recruitment patterns are not observed for yellowfin tuna in the EPO.” --- SAC11 INF-F

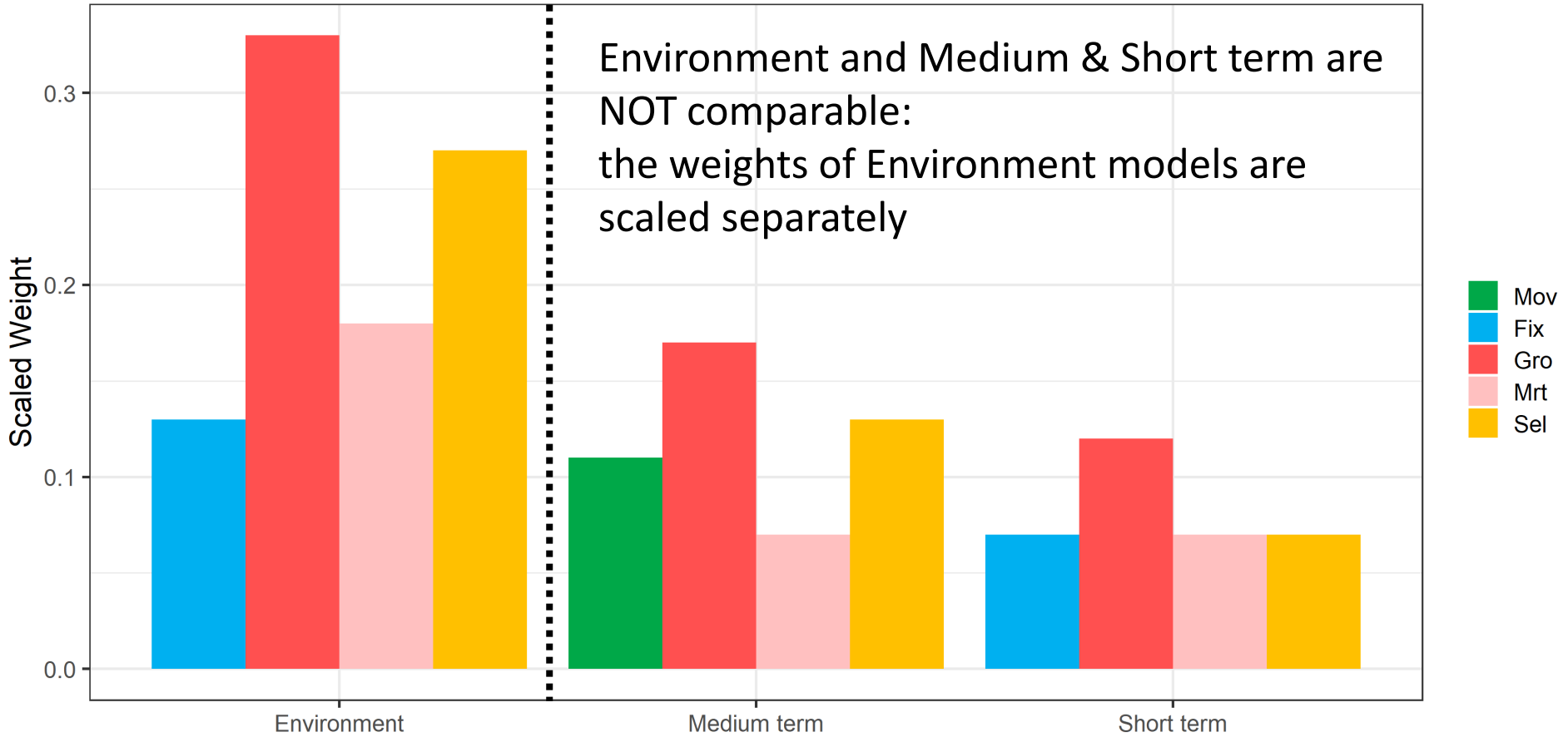
W(expert) – level 2

The weight of each hypothesis by each expert *a priori:*

- Weights are developed independently for levels 2A and 2B
- Joint weight computed

W(expert) – level 2

Experts favor estimating growth and dome-shaped selectivity



W(convergence) – level 2

Whether the model converged with a small maximum gradient and positive definite Hessian:

- All converged models have equal weight
- All none-converged models (NA in the table) have zero weight

Maximum gradient (e.g., 2.E-04 means 0.0002)

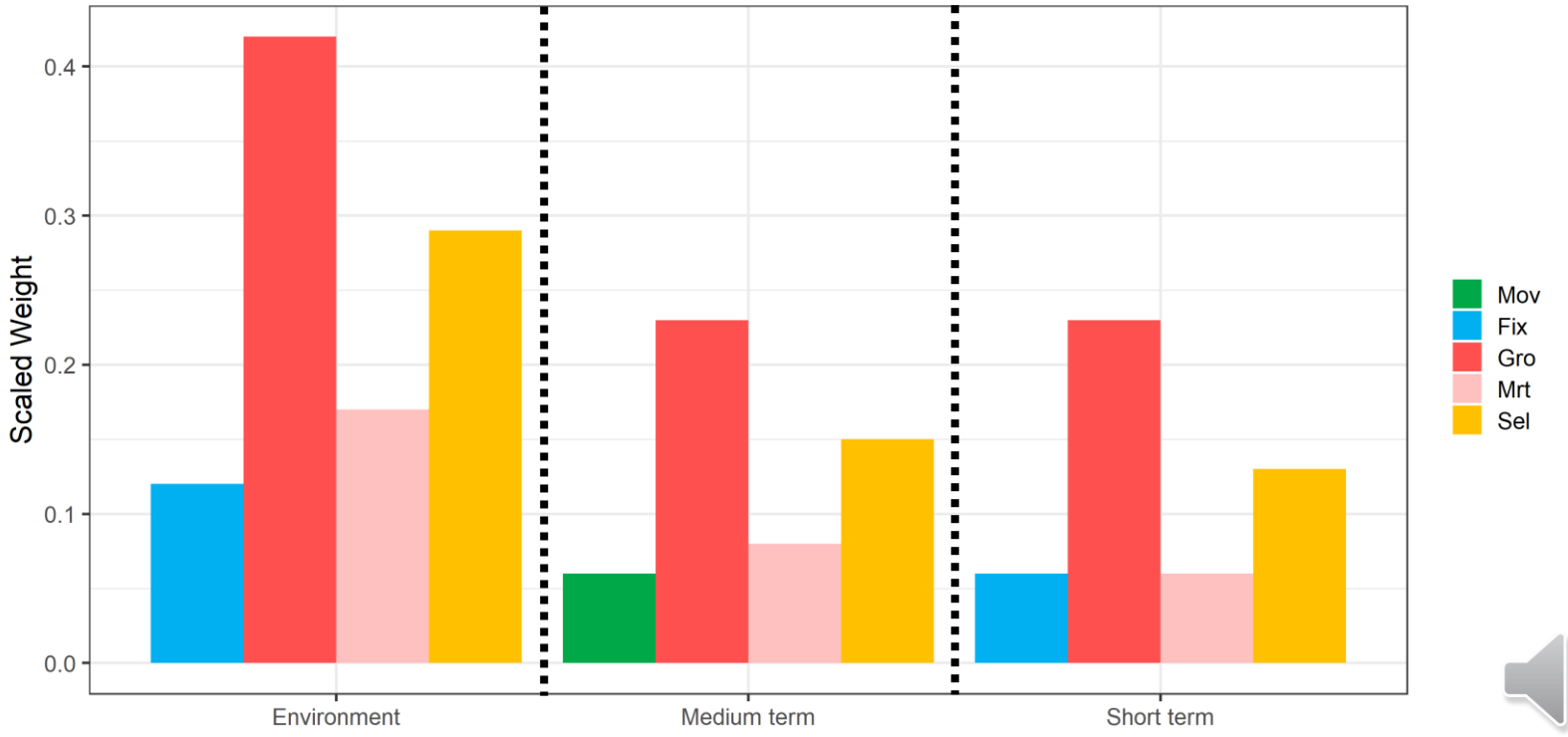
Steepness	Env-Fix	Env-Gro	Env-Mrt	Env-Sel	Gro	Mov	Mrt	Sel	Srt-Fix	Srt-Gro	Srt-Mrt	Srt-Sel
h = 1	2.E-04	9.E-05	4.E-05	7.E-05	3.E-05	7.E-05	6.E-05	1.E-03	5.E-05	8.E-05	8.E-05	7.E-05
h = 0.9	NA	3.E-05	2.E-04	6.E-05	8.E-05	1.E-04	6.E-05	9.E-04	7.E-05	3.E-04	5.E-05	2.E-05
h = 0.8	NA	8.E-05	5.E-05	7.E-05	5.E-05	1.E-04	1.E-03	1.E-03	6.E-03	1.E-03	1.E-02	4.E-05
h = 0.7	NA	2.E-03	7.E-05	5.E-05	1.E-04	2.E-04	7.E-05	2.E-04	NA	2.E-03	3.E-03	1.E-05

W(fit) – level 2

The support of data to each hypothesis:

- Approximation: Akaike Information Criterion (AIC)
- Linear weight from worst (0.25) to best (1) models based on AIC
- Short term models are weighted separately

Estimating growth fits best to data

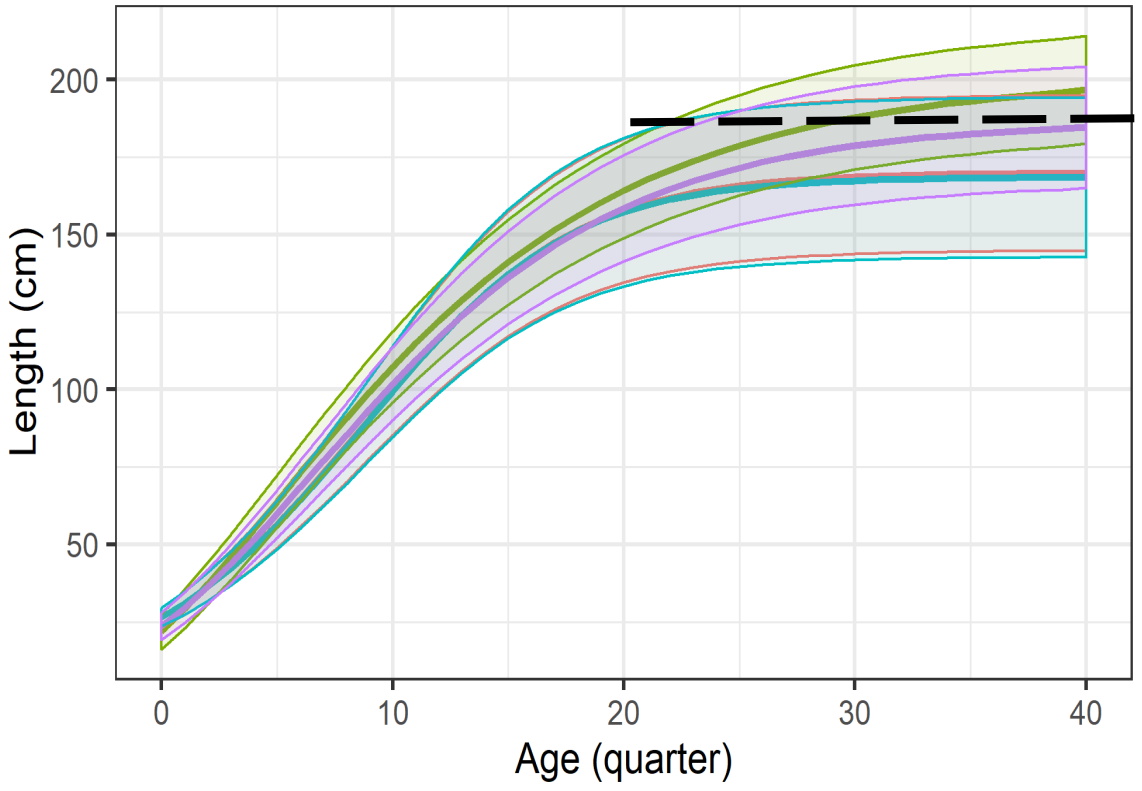


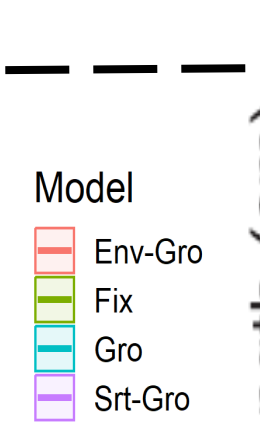
W(plausible parameters) – level 2

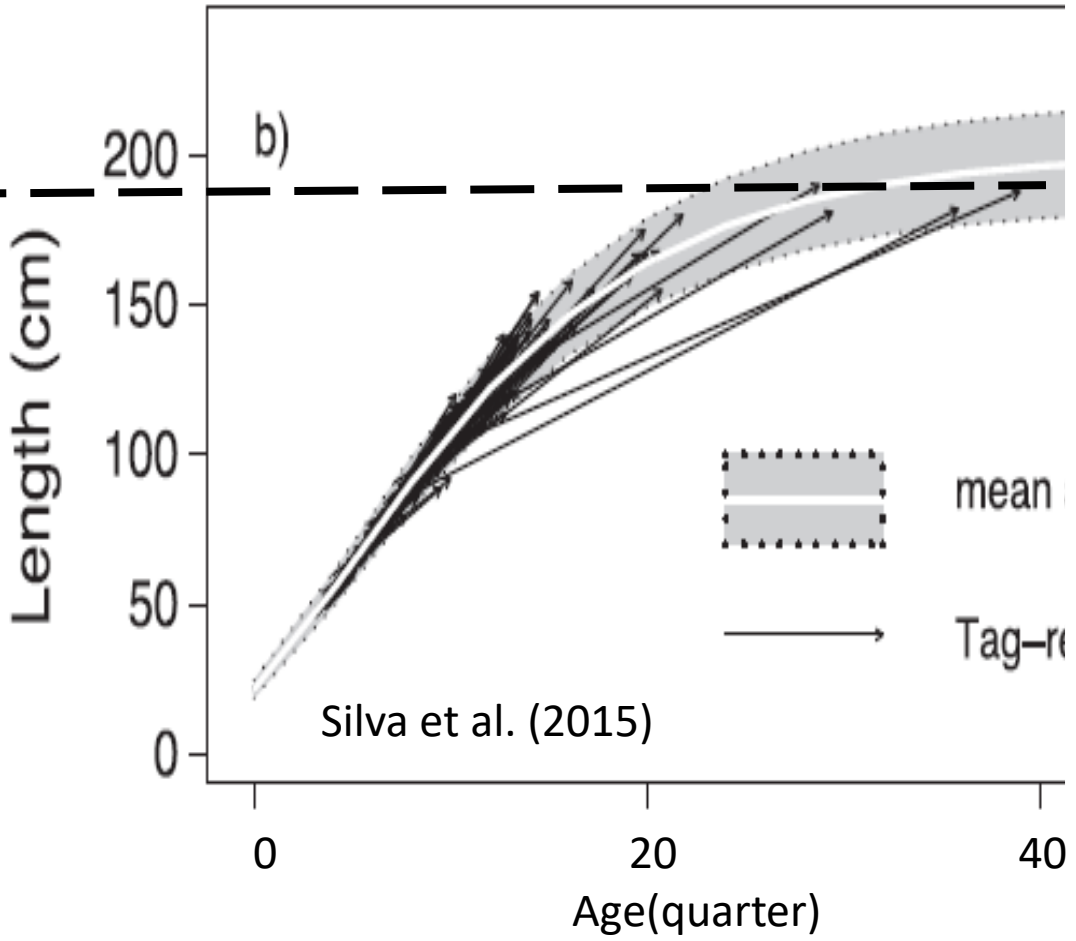
The realism of the parameters that represent hypotheses:

Are the parameters realistic compared to expert judgement, theory, other data not used in the model?

W(plausible parameters): growth

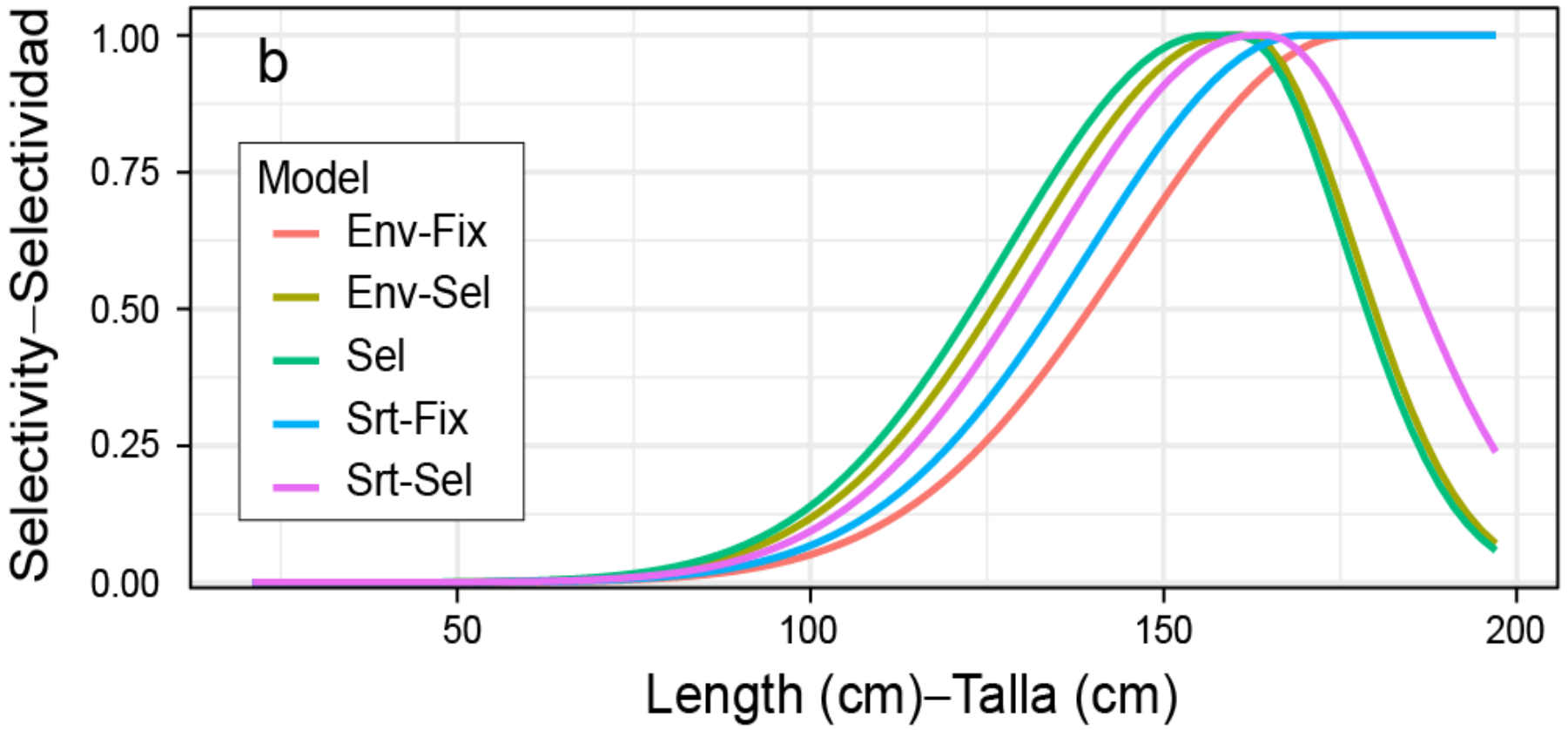




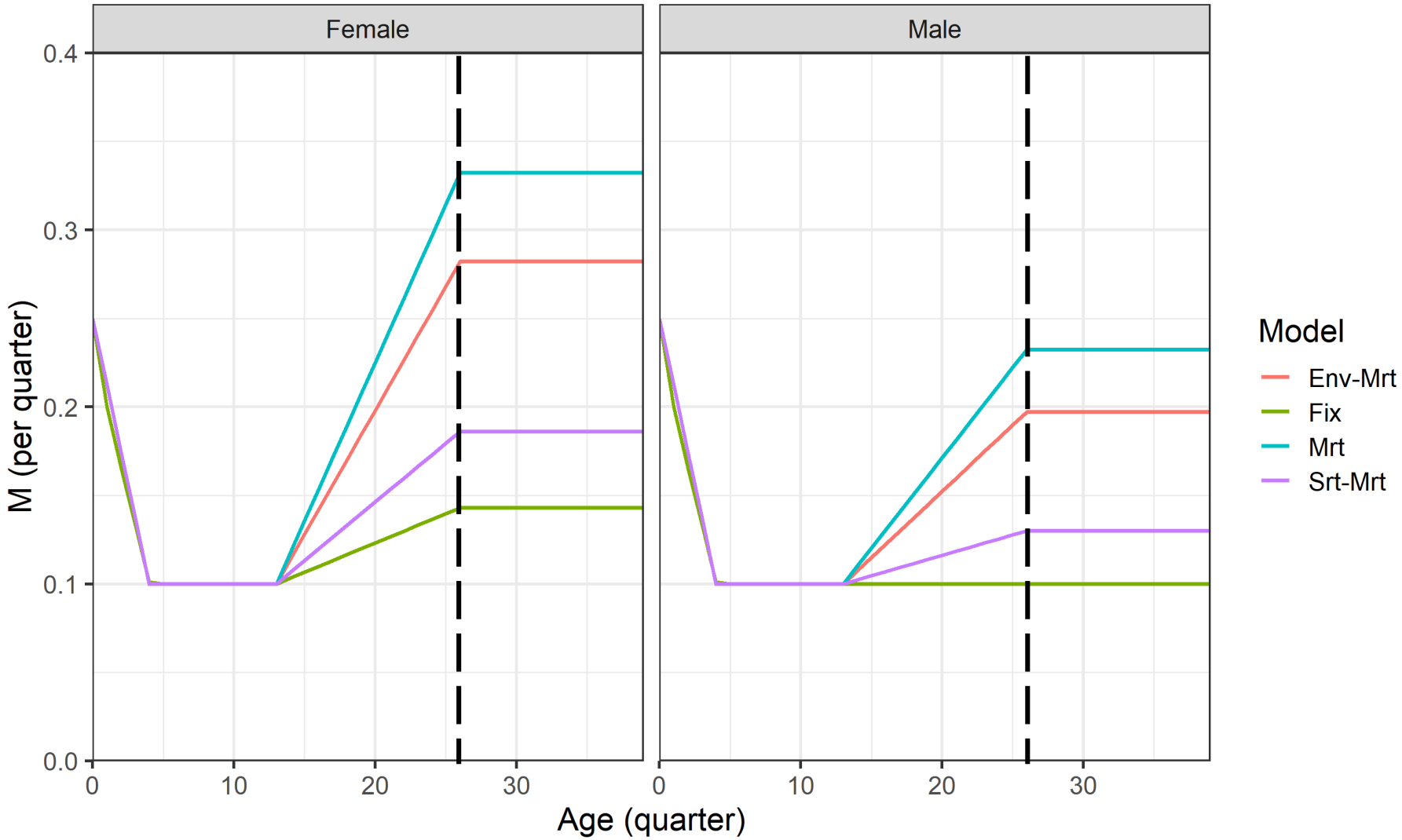


W(plausible parameters): selectivity of Fishery 2

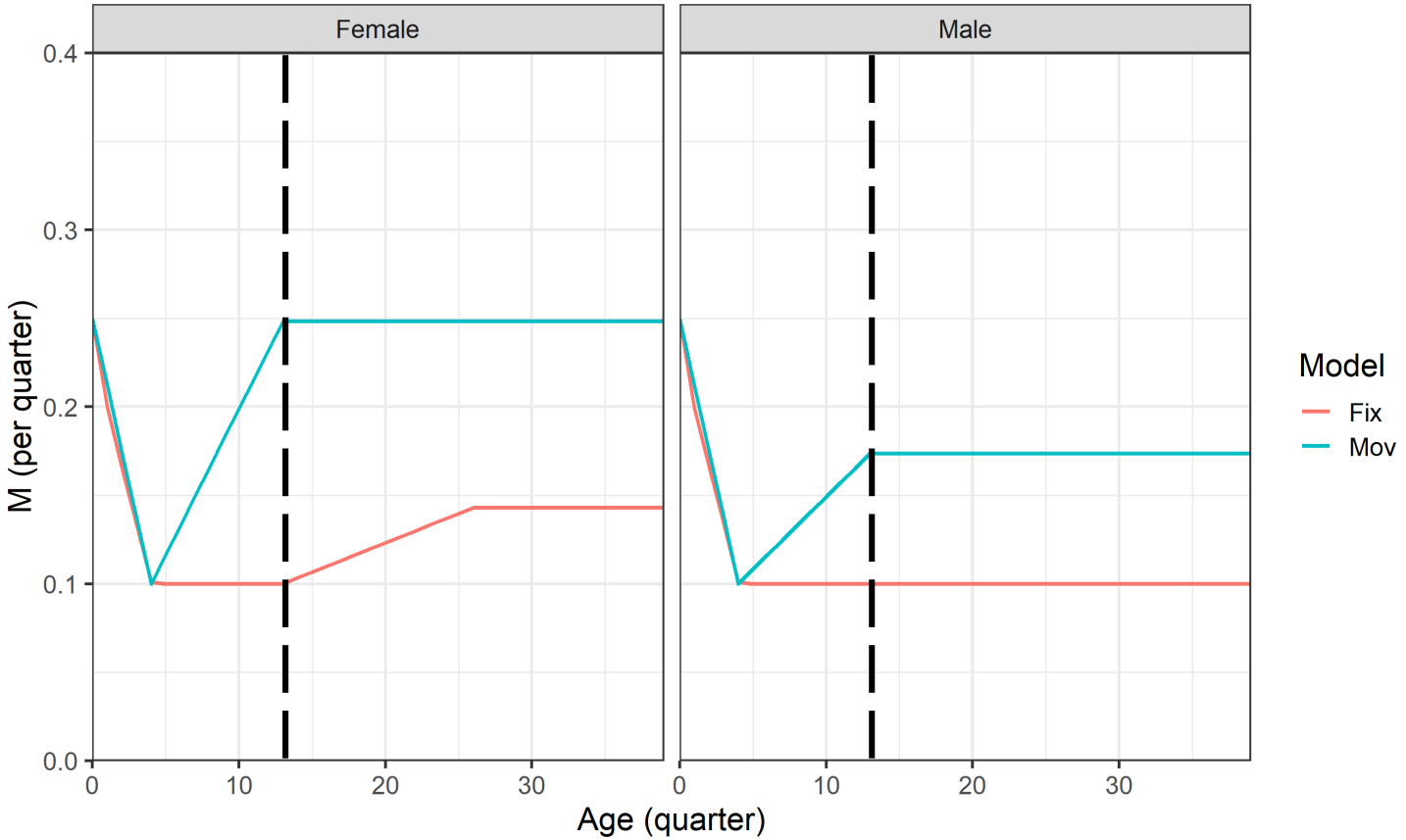
Fishery 2 (A2-LL-n): the longline fishery that catches the highest proportional of very large bigeye



W(plausible parameters): natural mortality



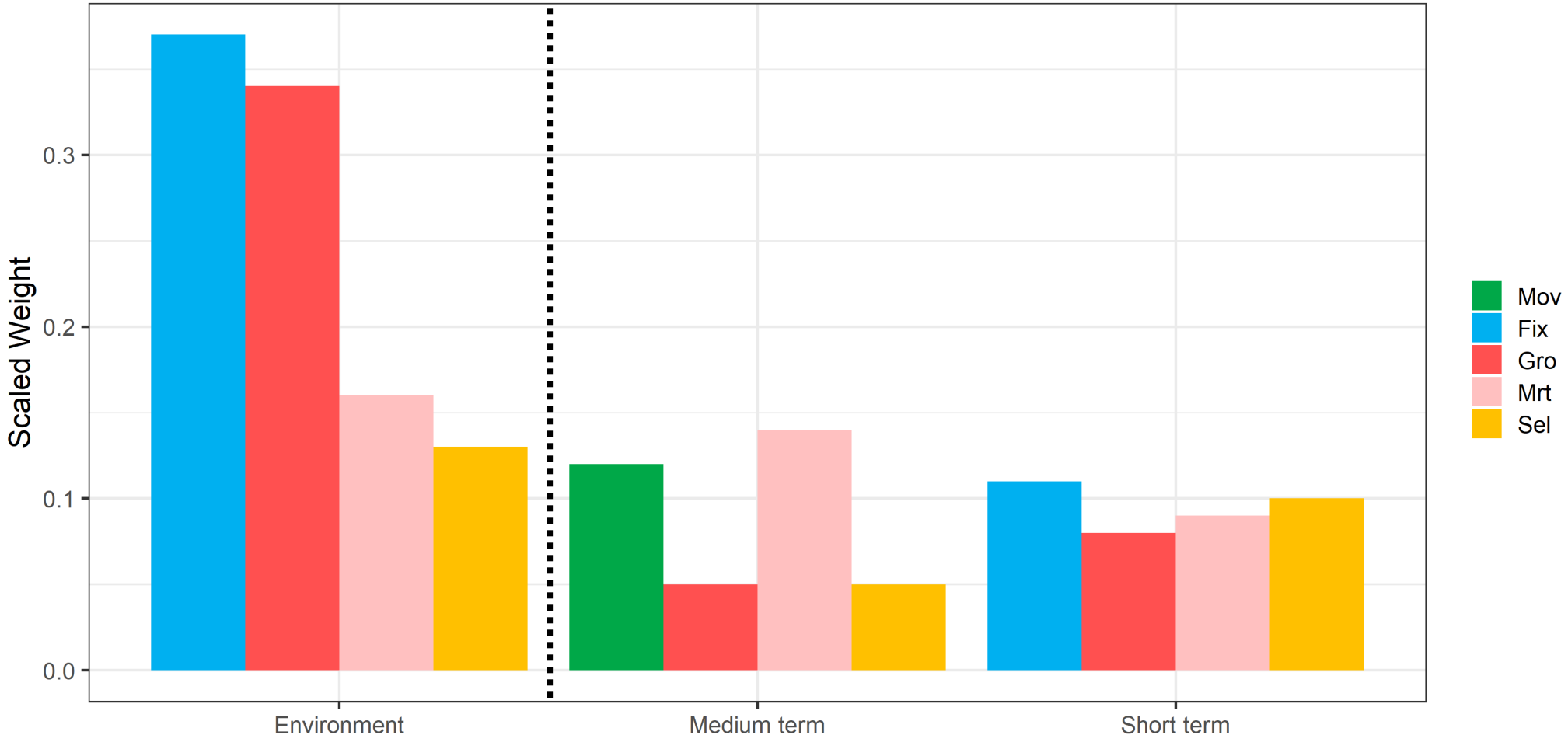
W(plausible parameters): movement



Natural mortality is used to represent the pre-adult movement between the EPO and WCPO

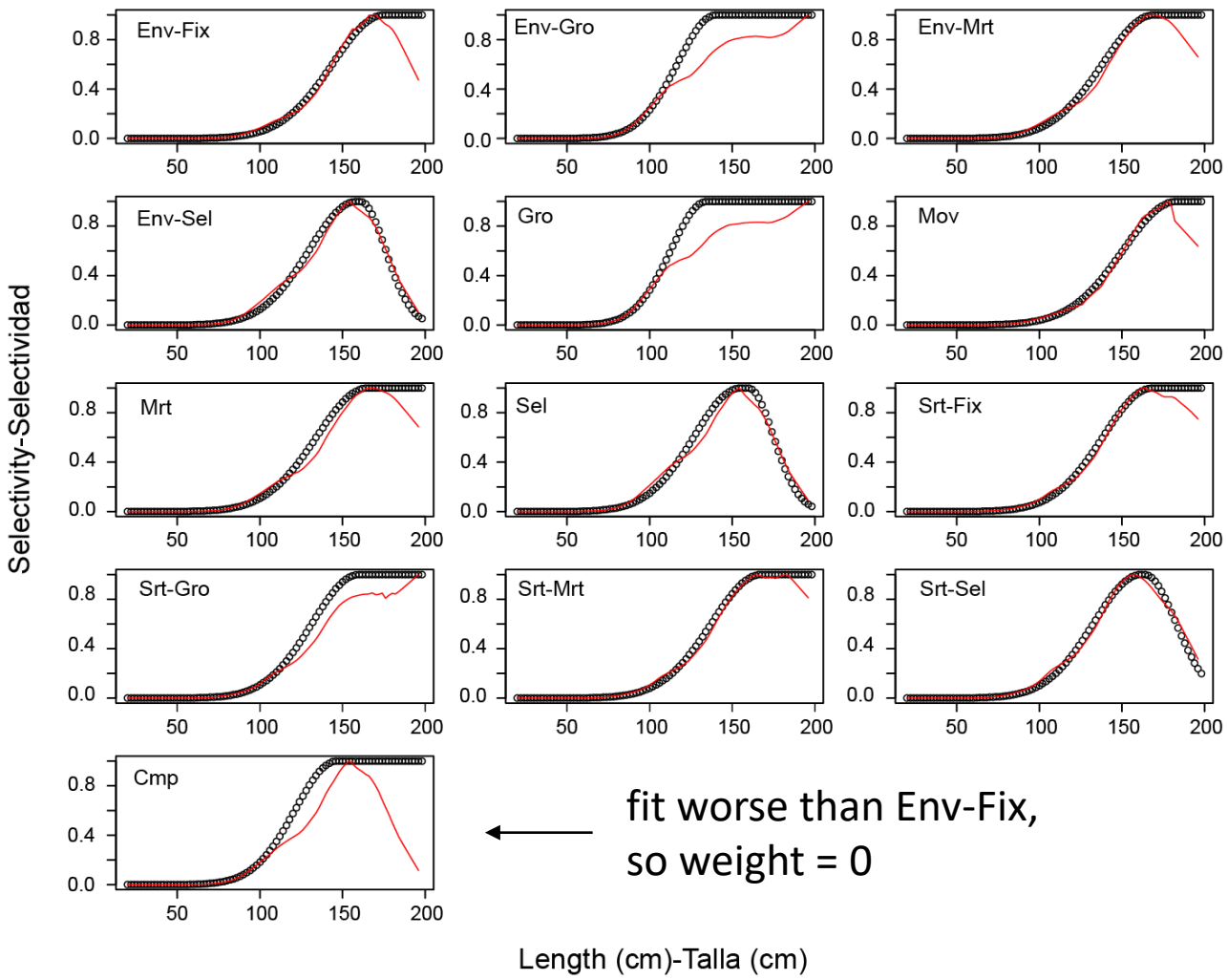
Higher natural mortality: moving from the EPO to WCPO
Lower natural mortality: moving from the WCPO to EPO

W(plausible parameters) – level 2



W("empirical" selectivity) – level 2

How well the estimated selectivity curve represents the empirical selectivity for Fishery 2?



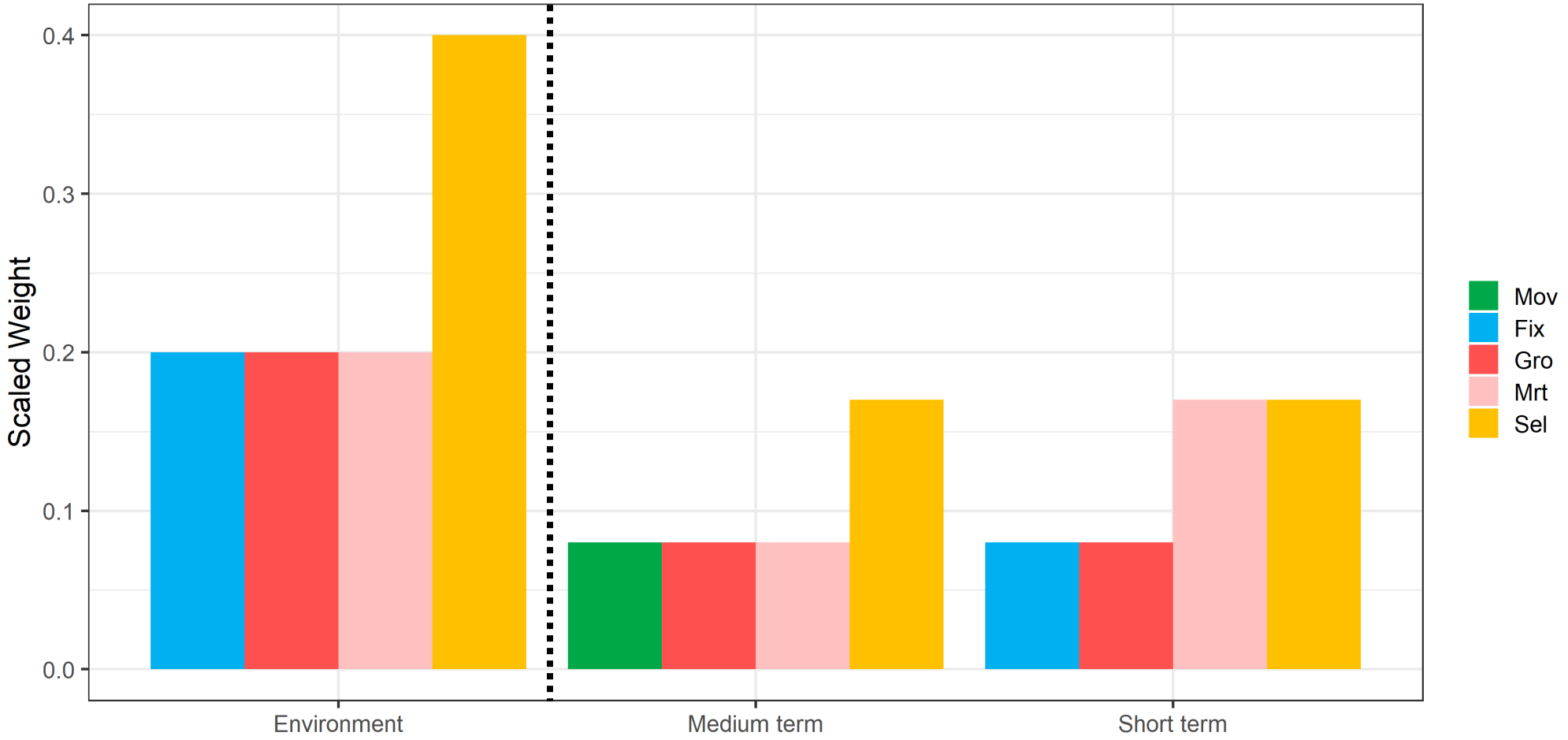
fit worse than Env-Fix, so weight = 0

red line: "empirical" selectivity of Fishery 2 since 2000

black dots: estimated selectivity of Fishery 2 since 2000

W("empirical" selectivity) – level 2

dome-shaped selectivity: high weight



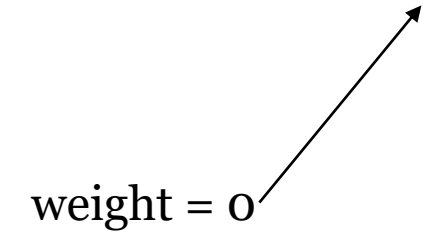
W(recruitment shift) – level 2

Note: The metric is not applicable to the environment and short term models, so they all have high weight

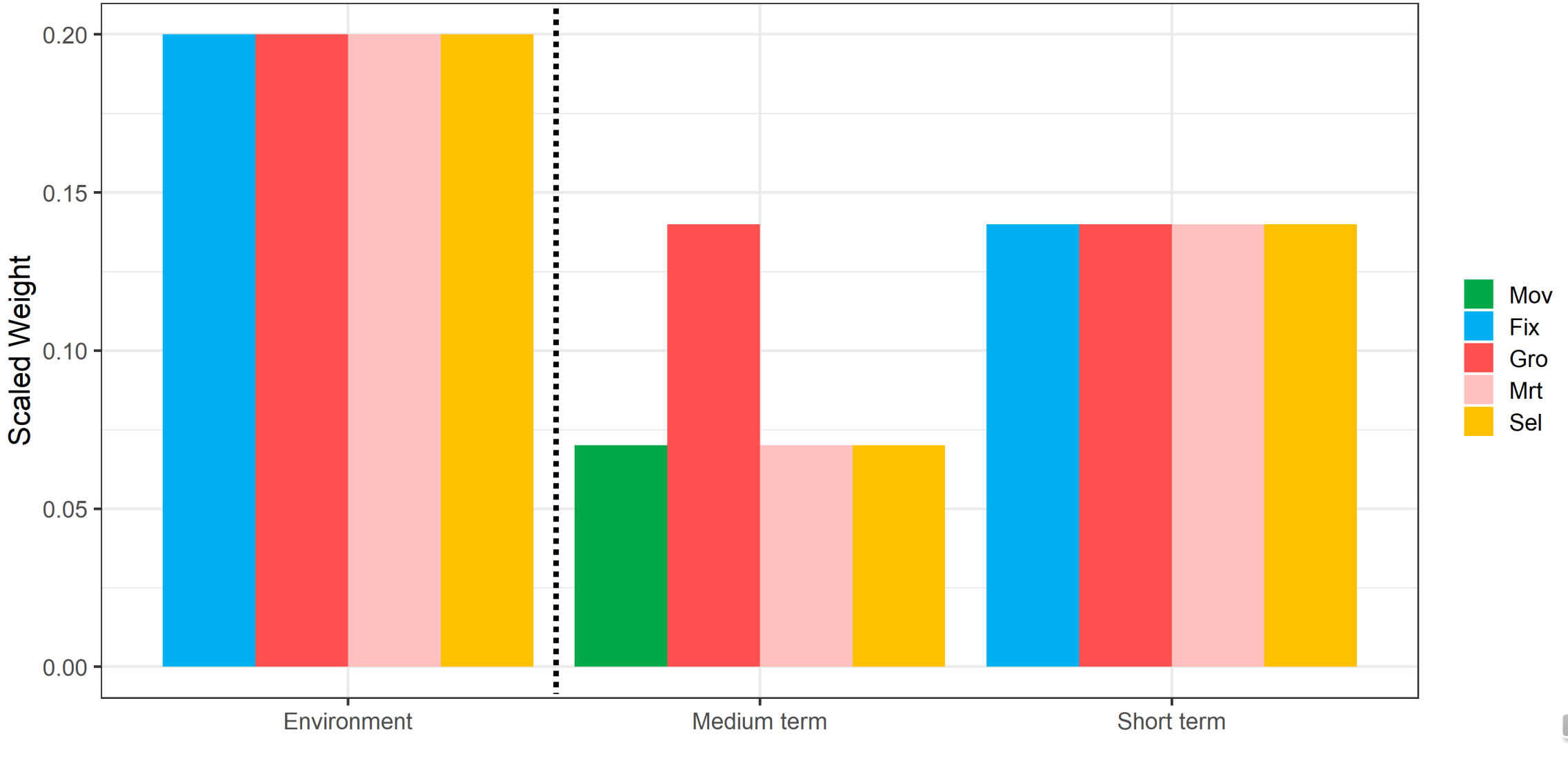
Regime Shift category	Weight category
1.75 < Rshift	None
1.50 < Rshift ≤ 1.75	Low
1.25 < Rshift ≤ 1.5	Medium
Rshift ≤ 1.25	High
Regime shift is real	NA

The ratio of the median recruitment in 1994-2019 to that in 1979-1993

Gro	**Mov**	**Mrt**	**Sel**	**Ind**
1.2	1.3	1.4	1.3	2.1



W(recruitment shift) – level 2

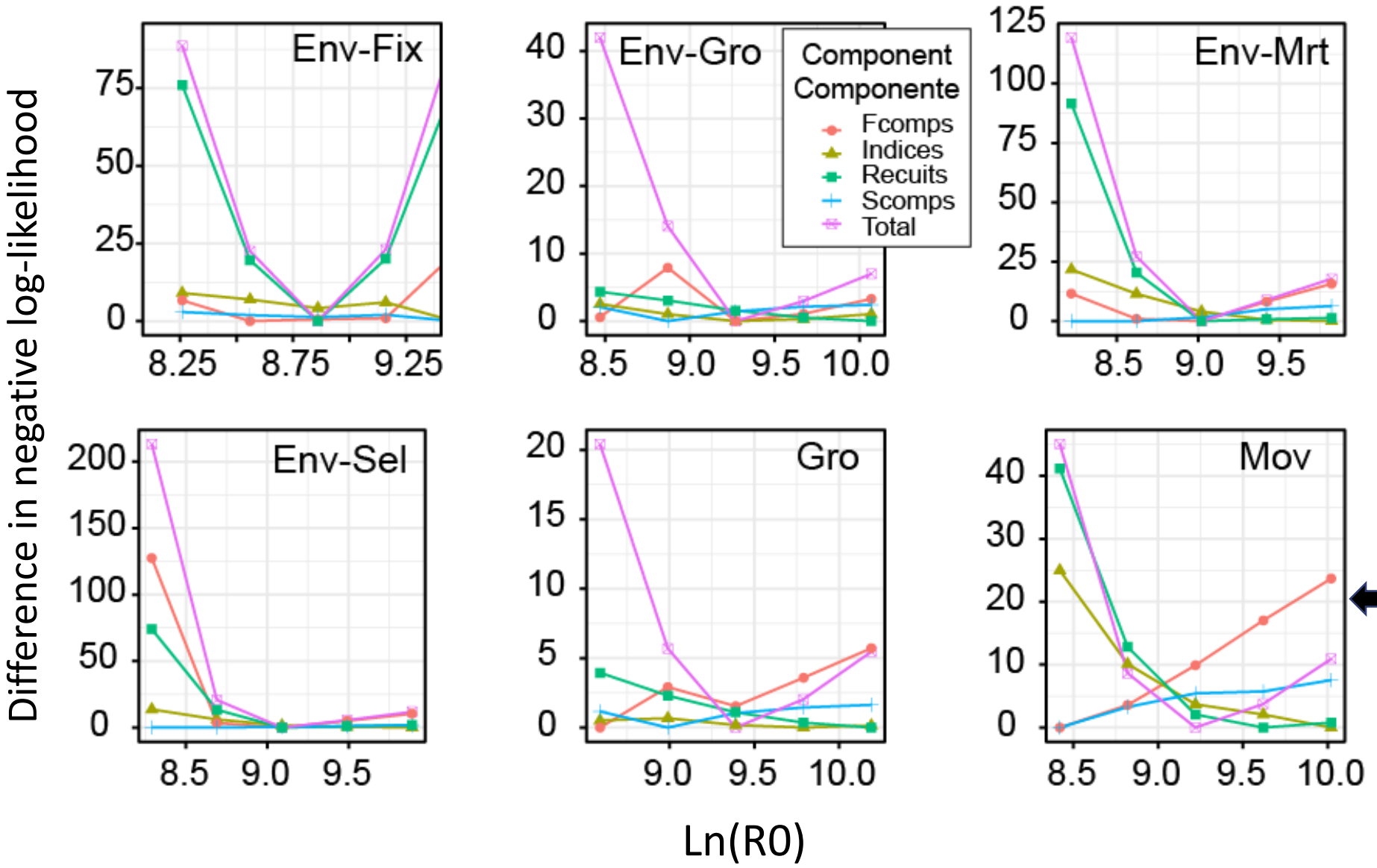


W(diagnostics) – level 2

How well do models perform according to different diagnostics:

- Age-structured production model and R_0 profile
- Retrospective analysis
- Residual patterns

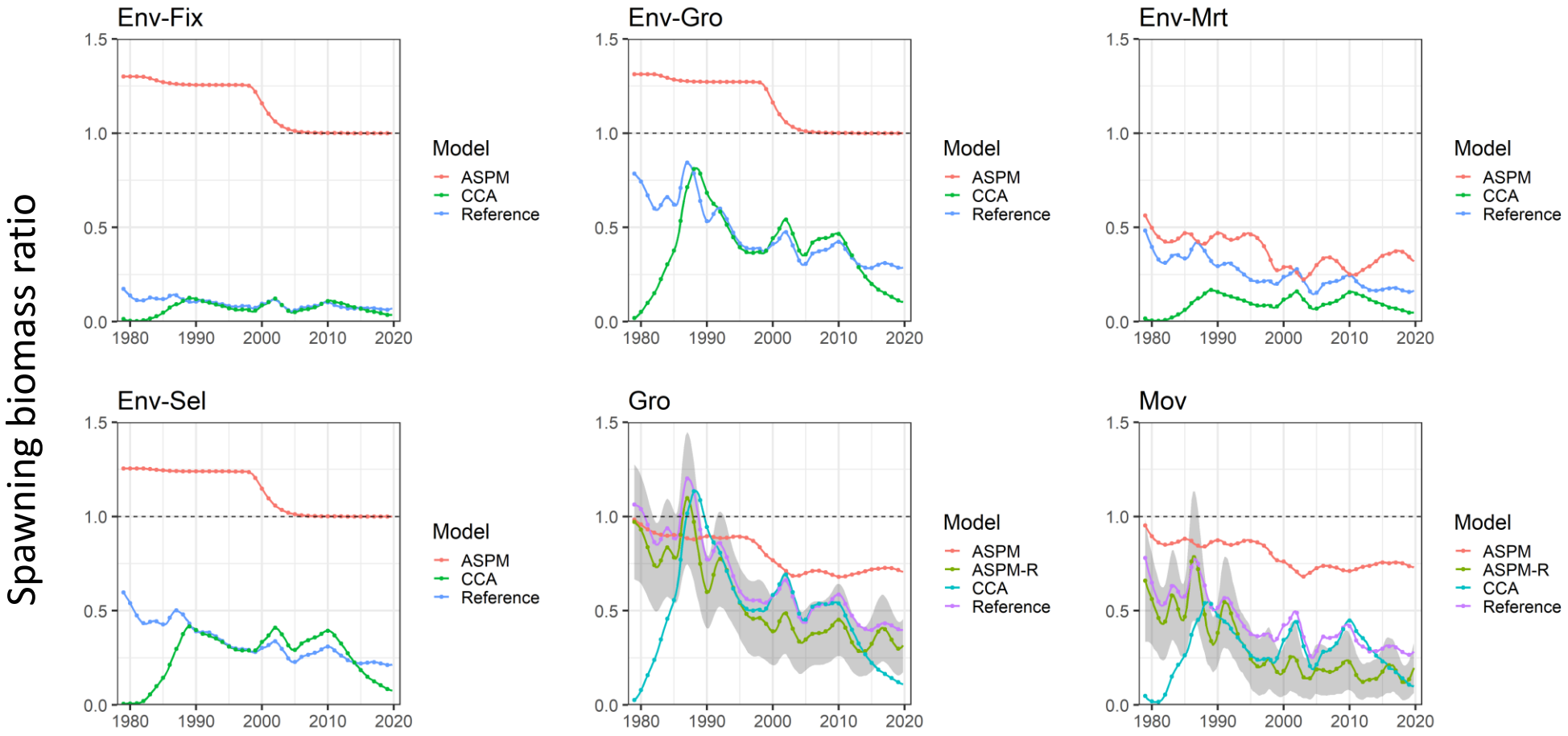
W(ASPM and R0 profile) – level 2



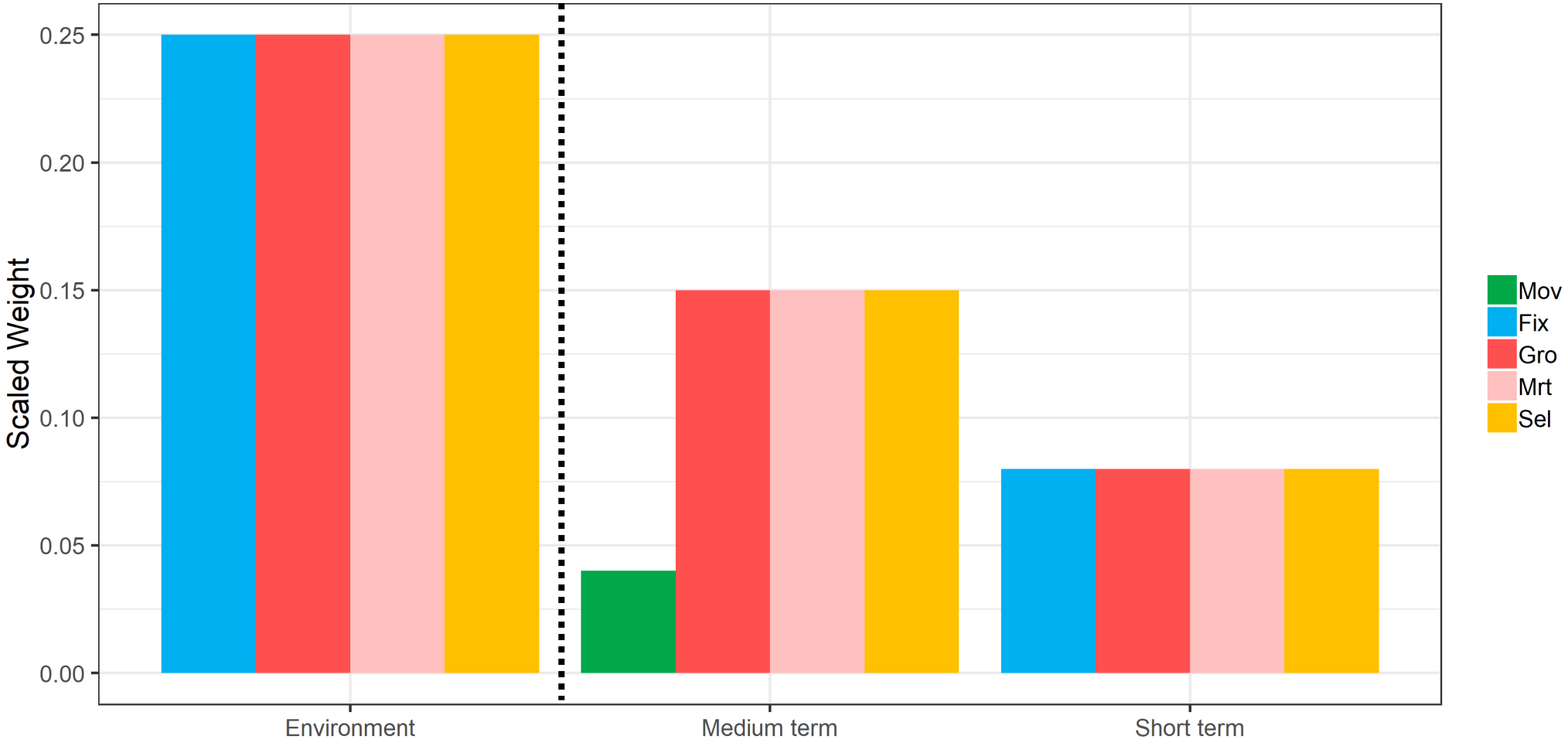
Contradictory information provided by Fishery compositions and indices of abundance

W(ASPM and R0 profile) – level 2

No ASPM-R: the model does not converge, which is therefore assumed to have large confidence interval



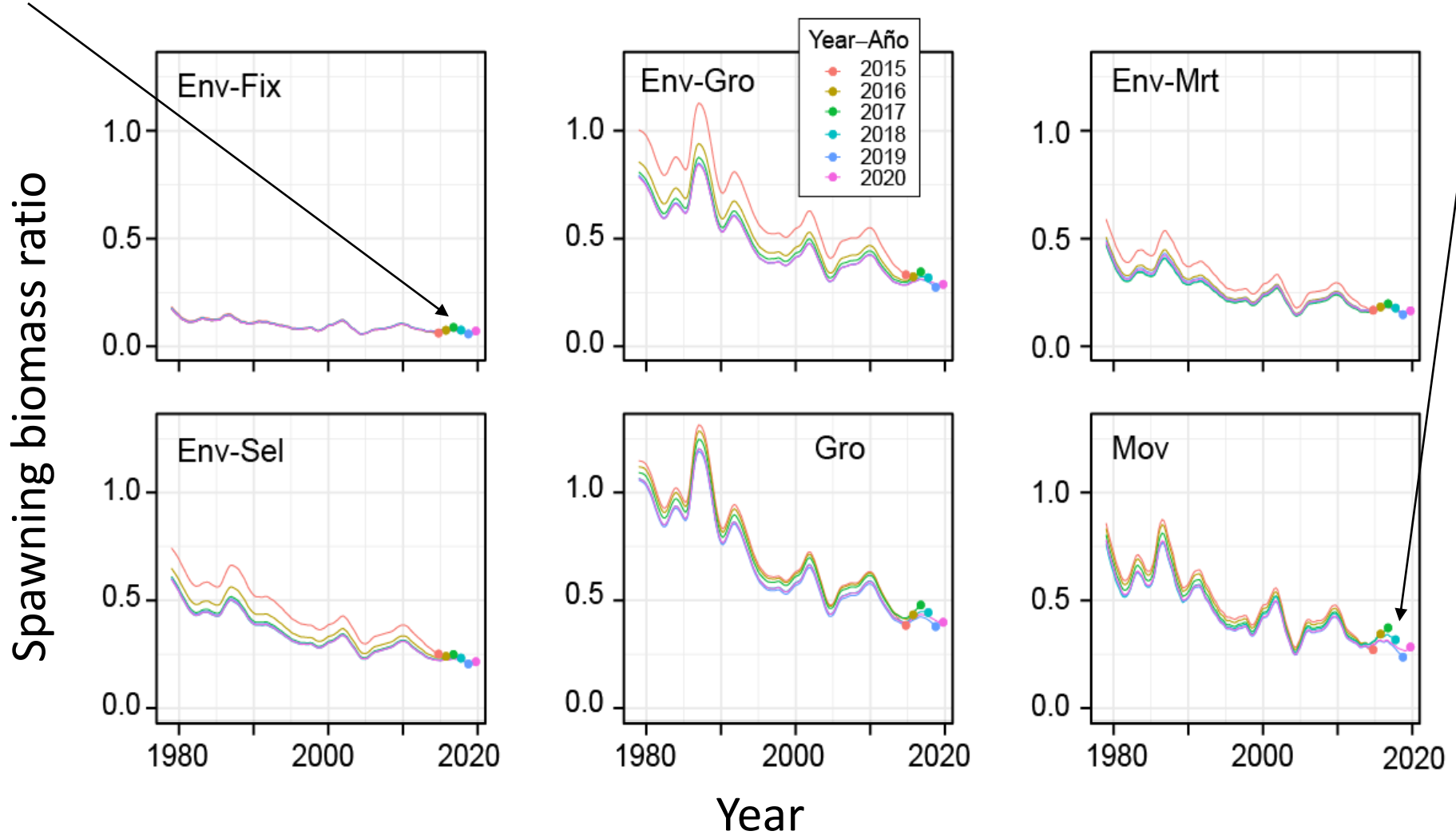
W(ASPM and R0 profile) – level 2



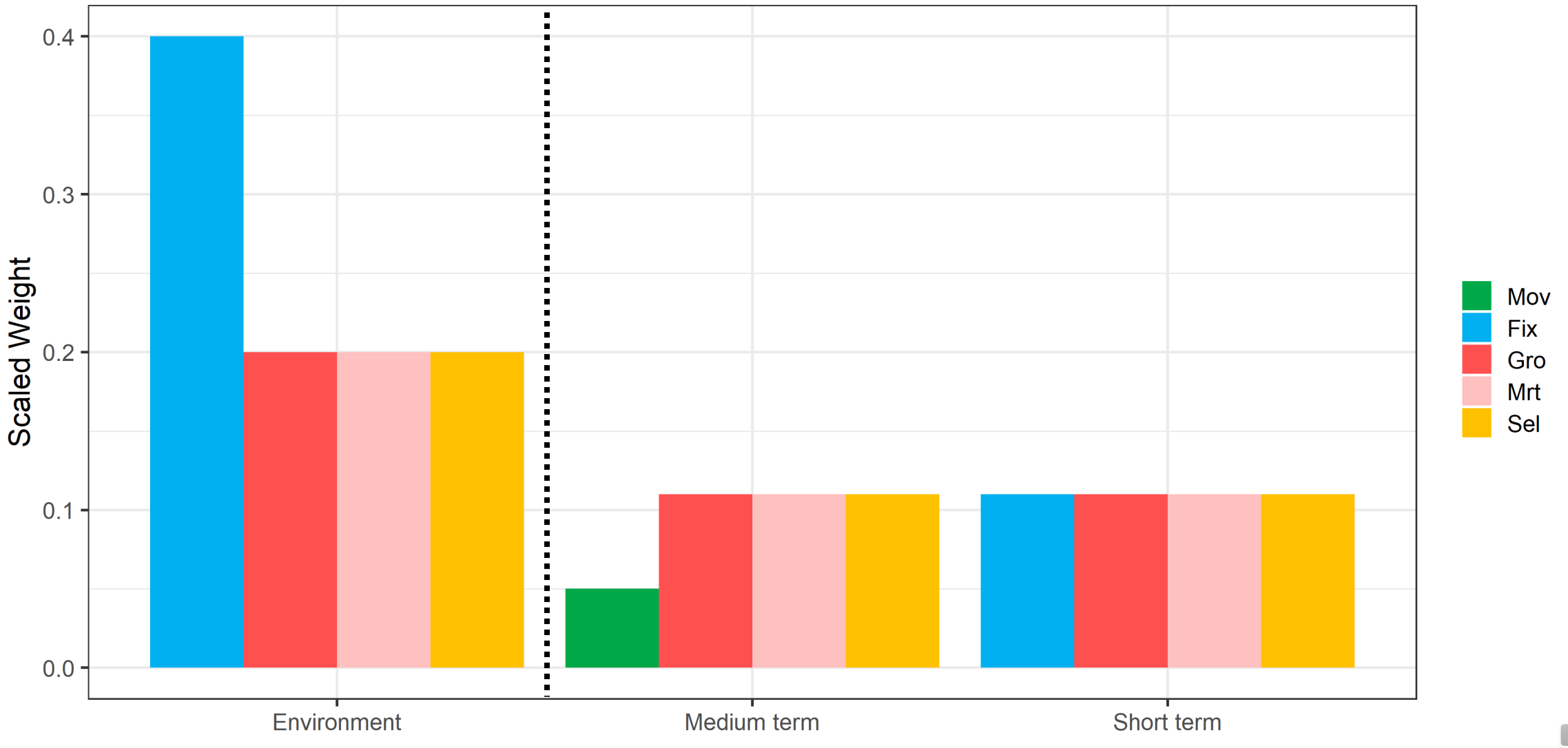
W(retrospective) – level 2

small retrospective pattern: high weight

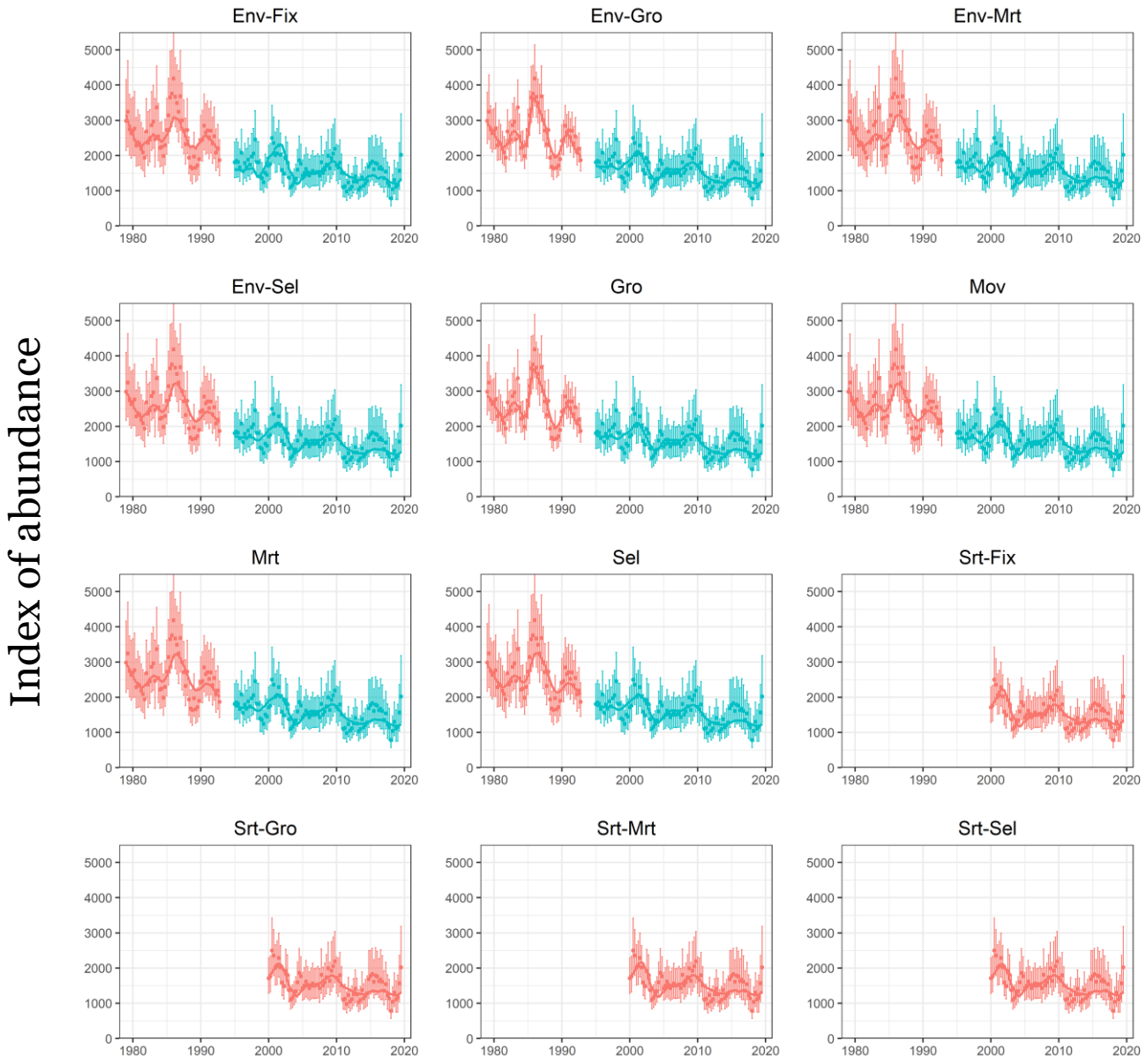
large retrospective pattern: low weight



W(retrospective) – level 2



W(residual pattern) – level 2



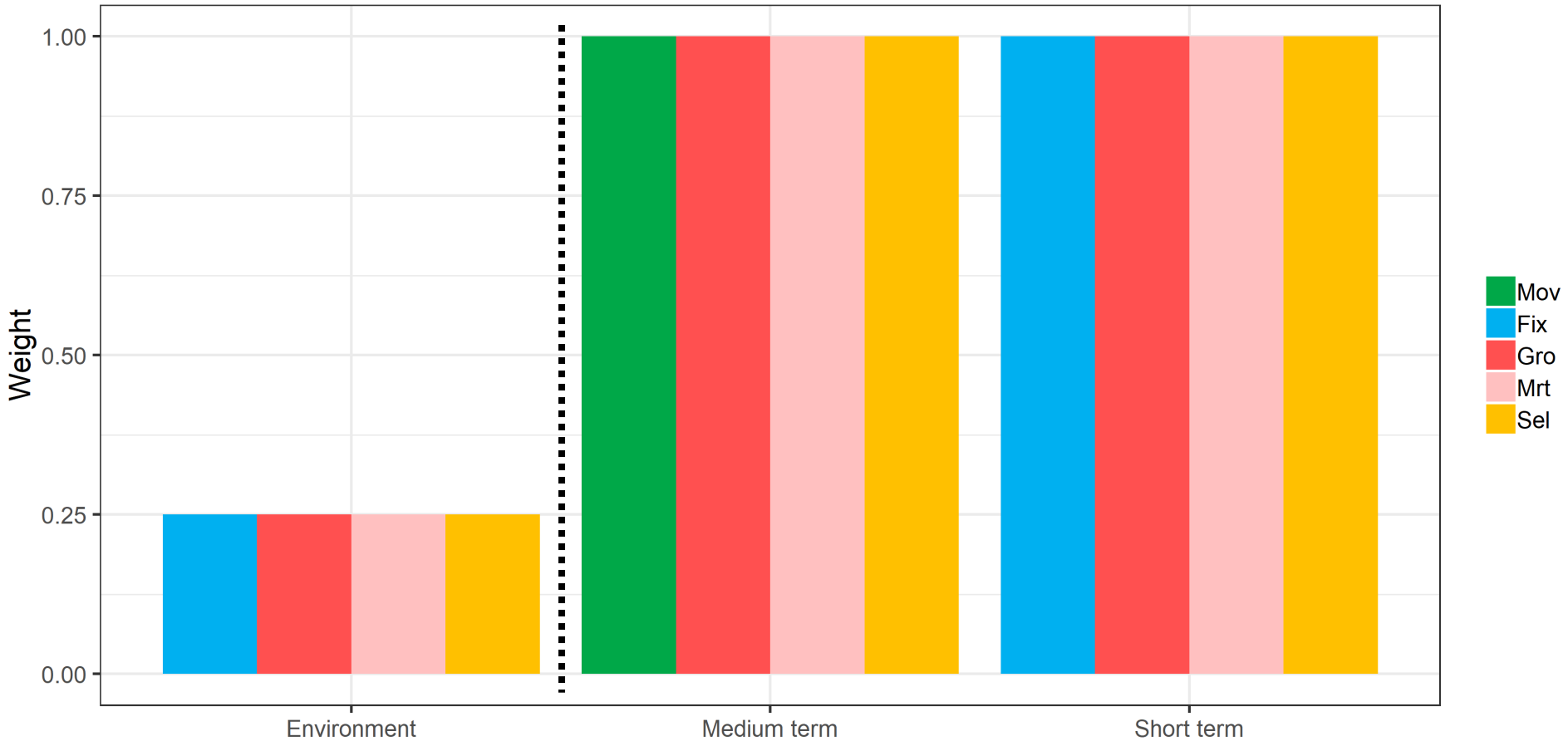
Very similar residual patterns among models, so all models have equal W(residual pattern)

Combined weights – level 2

The combination of the weights in each category:

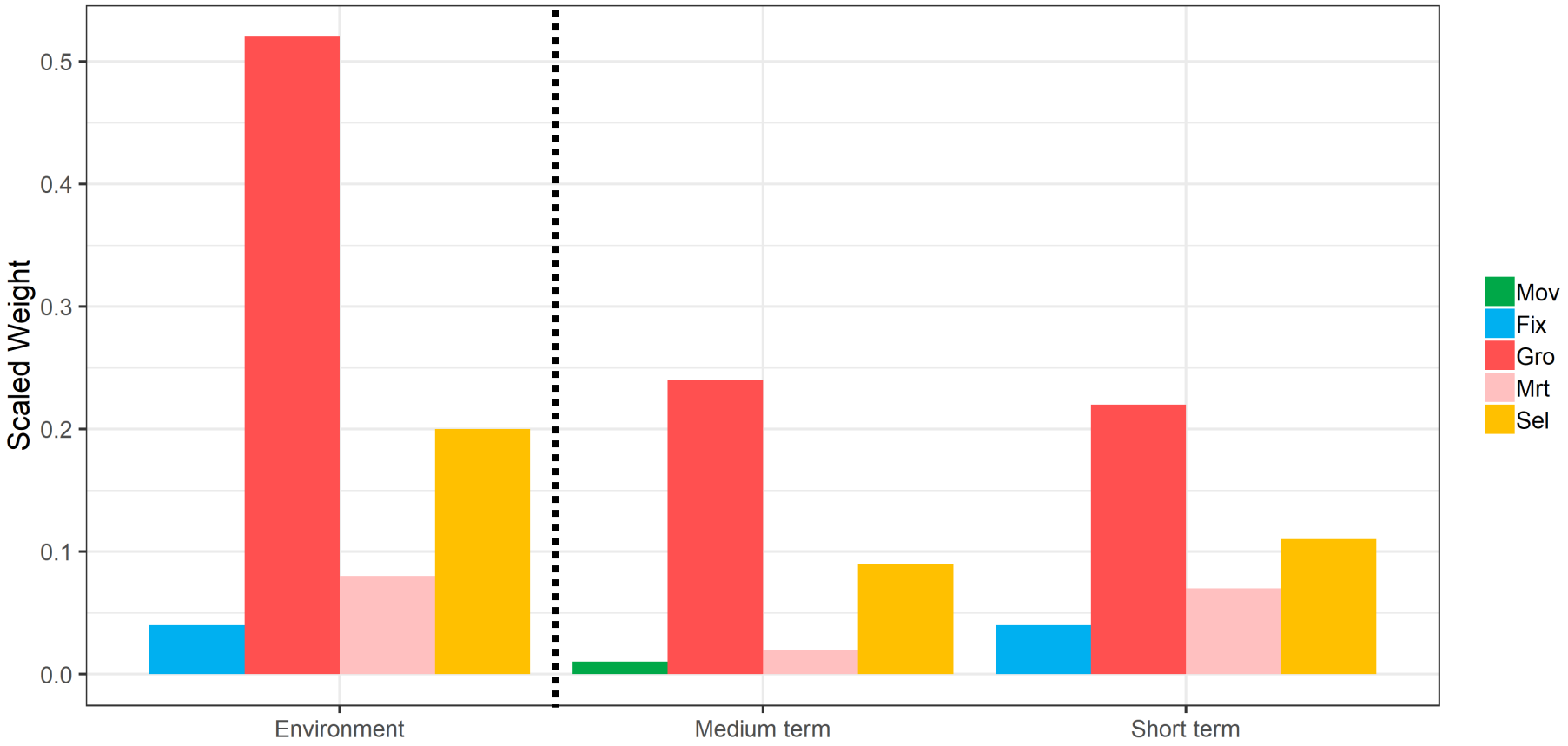
- The weights for individual diagnostics are summed to create an overall diagnostics metric
- Weights are multiplied
- Weights are rescaled to sum to 1

Level1 Final weights



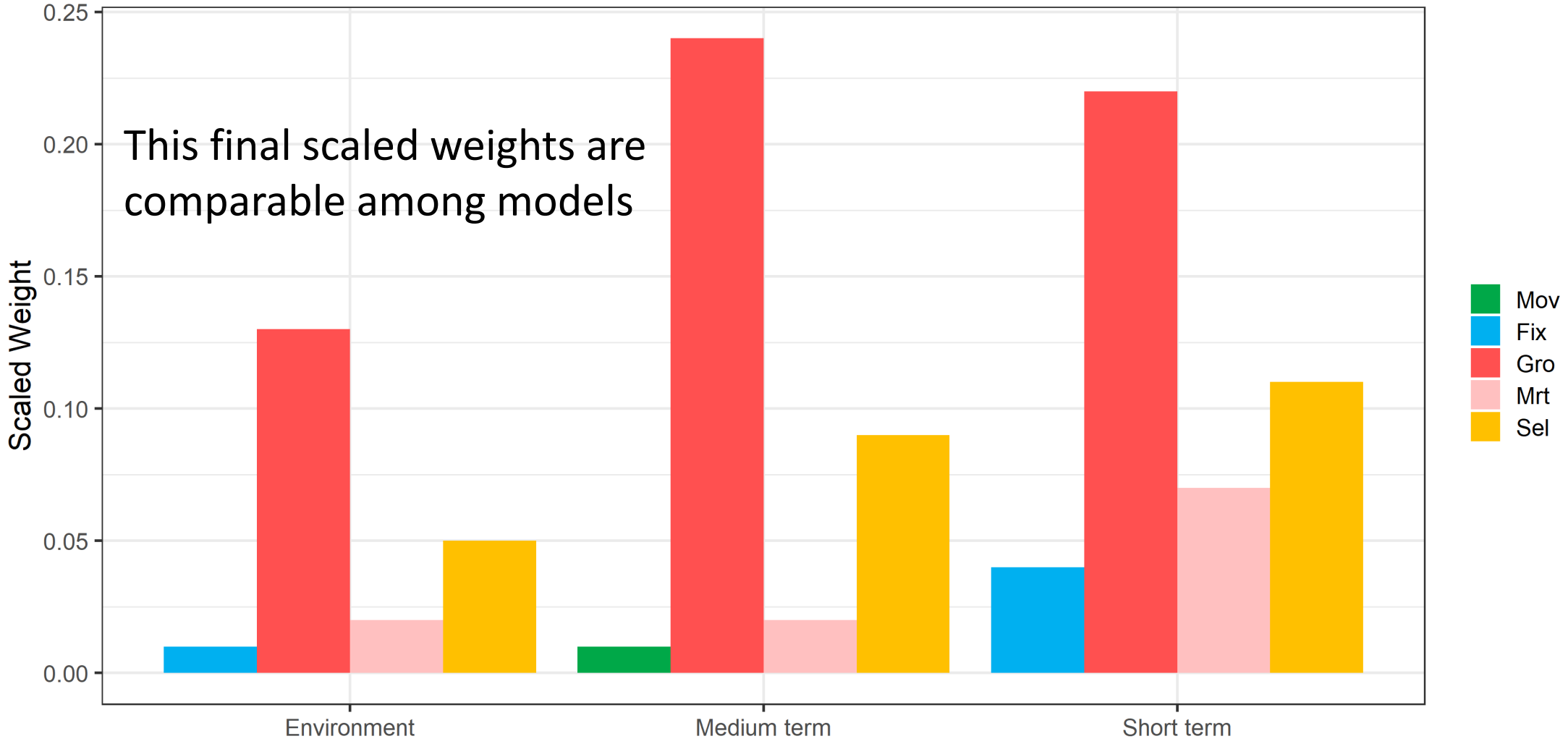
- High weight for recruitment regime shift being not real
- Low weight for recruitment regime shift being real

Level2 Final weights



- High weight for estimating growth and dome-shaped selectivity

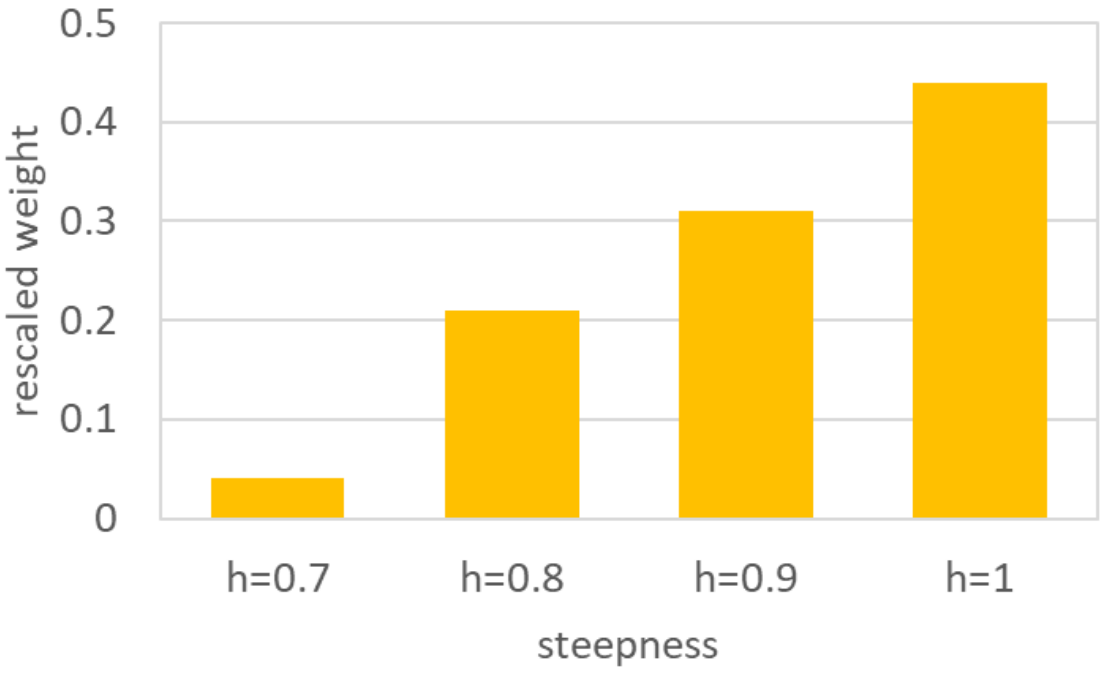
Level1 * Level2 Final scaled weights



W(steepness) – Level 3

The weight given to different steepness values:

- Weighted by every expert considering evidence regarding steepness
- Weights are combined across experts



Conclusions

- Model weights are necessary to combine model results and estimate probabilities of exceeding reference points
- The approach developed by the staff allows for a systematic review of several aspects of model performance
- This novel approach is more appropriate than un-weighted model averaging

Next step in the risk analysis approach

Described in Maunder et al. 2020 (SAC-11- INF-F):

1. **Identify alternative hypotheses** (‘*states of nature*’) about the population dynamics of the stock that address the main issues in the assessments
   - YFT: SAC-11-J; BET: SAC-11 INF-F
2. **Implement stock assessment models representing alternative hypotheses**
   - YFT: SAC-11-07; BET: SAC-11-06
3. **Assign relative weights to each hypothesis (model)**
   - YFT: SAC-11 INF-J; BET: SAC-11 INF-F
4. **Compute combined probability distributions for management quantities using model relative weights**
   - SAC-11-08

CIAT

IATTC

Thank you