Index Based Methods and Control Rules 2020 Research Track Assessment

Dr. Christopher M. Legault
Chair of the Index Based Methods and Control Rules Working Group

Presentation for
Mid-Atlantic Fishery Management Council's Scientific and Statistical Committee
9 March 2021

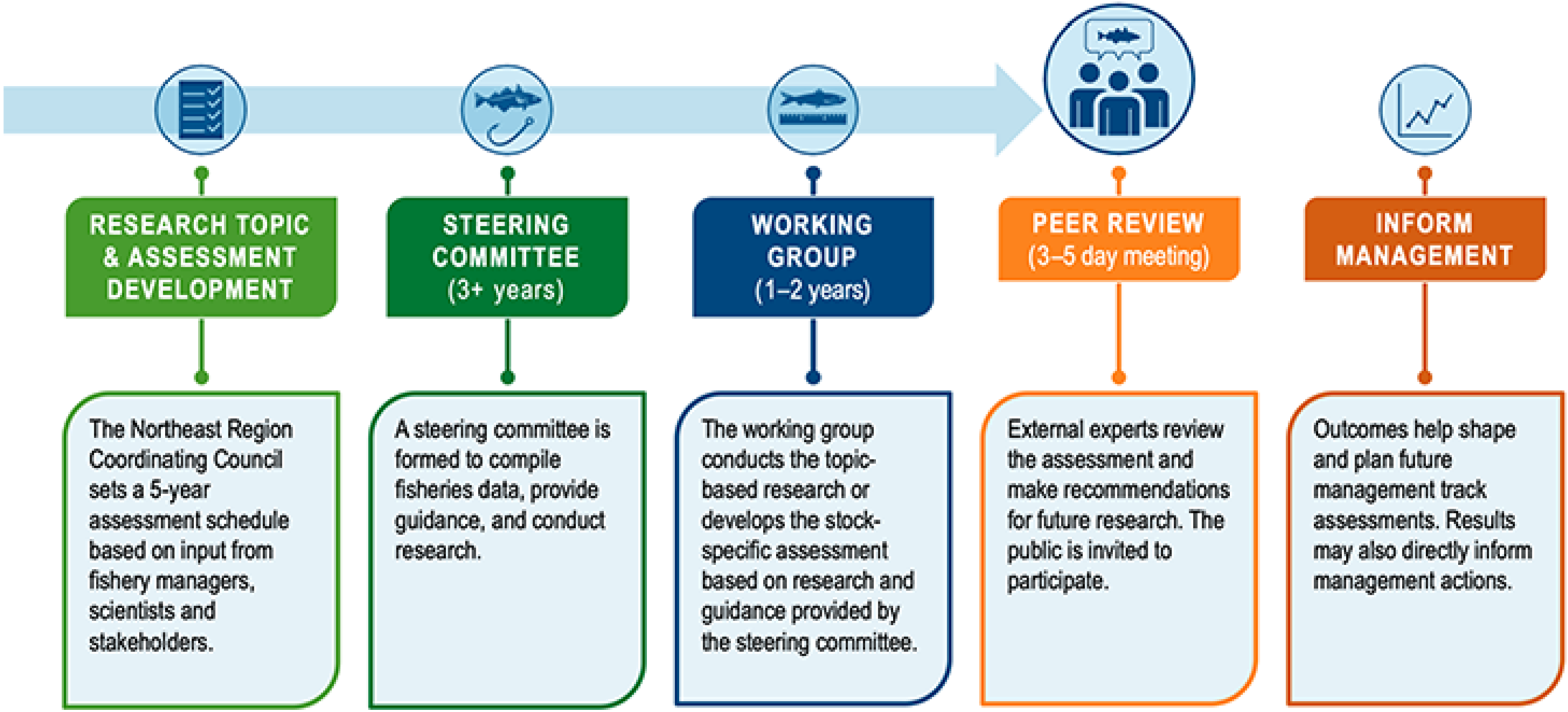
BLUF

- Developed framework
- Found two groups of IBMs
- No IBM better than SCAA

Genesis

- Retrospective patterns an issue for some (but not all) assessments in Northeast region
- Strong retrospectives led to 7 age-based models being rejected and replaced by index-based methods
- There are a range of index-based methods in the region
  - Both rejected age-based and always index-based
- Guidelines for picking an index-based method would help
- First Topic-based (instead of stock-specific) Research Track

RESEARCH TRACK STOCK ASSESSMENTS



https://s3.amazonaws.com/nefmc.org/Stock-assessment-process-June2020.pdf

Actual Timeline

- IBM Research Track approved 16 May 2019 by NRCC
- TORs finalized 10 Jan 2020
- WG formed 16 March 2020
- Weekly meetings began 26 March 2020
  - Total of 41 meetings
- Peer review 7-11 Dec 2020

10 months

8 months

https://www.fisheries.noaa.gov/new-england-mid-atlantic/population-assessments/stock-assessment-working-group-index-based-methods-and-control-rules

IBMWG

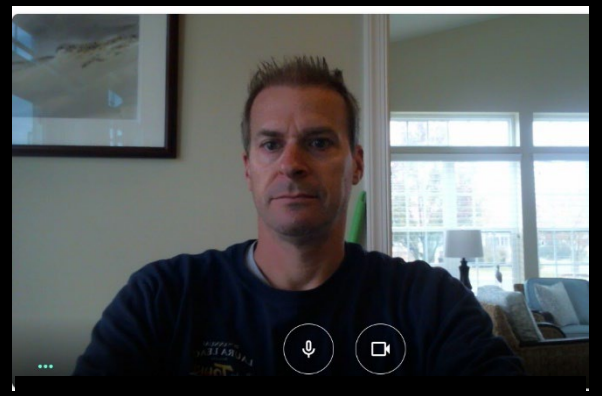

Andy Jones (NMFS)


John Wiedenmann (Rutgers)


Jon Deroba (NMFS)

HM =Honorary WG Member


Chris Legault (NMFS)


Brandon Muffley (MAFMC)


Gavin Fay (SMAST)


Liz Brooks (NMFS, HM)

Tim Miller (NMFS)


Rich Bell (TNC)


Jamie Cournane (NEFMC)

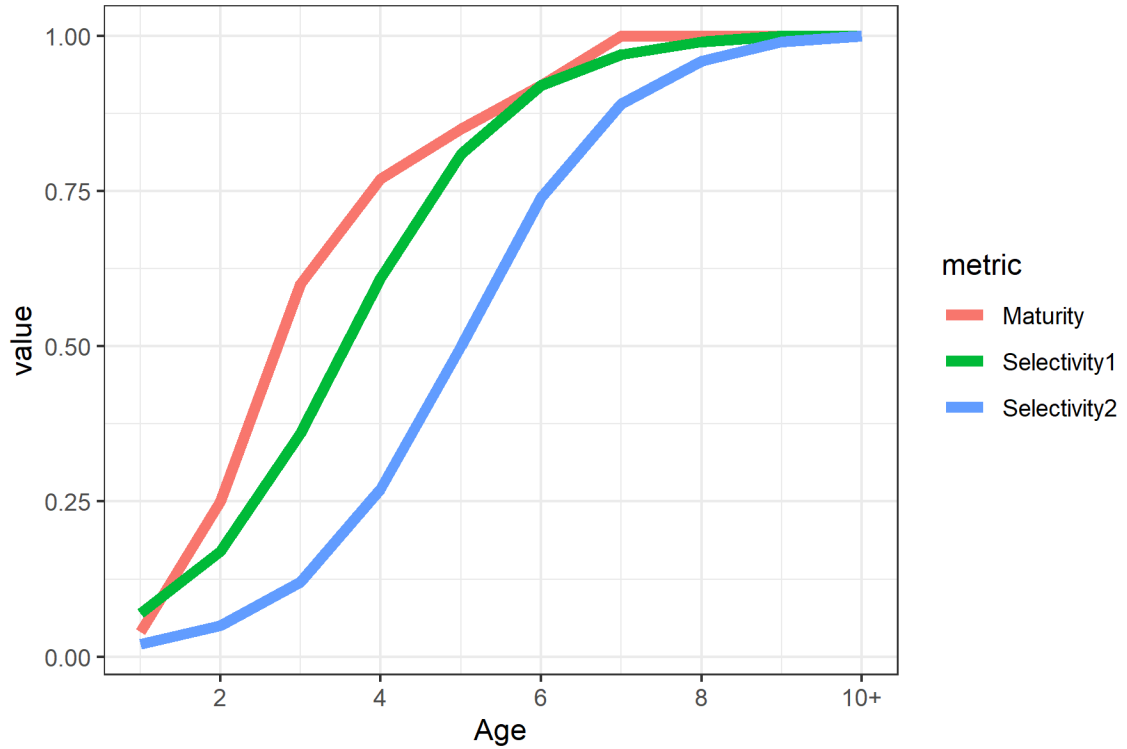

Joe Langan (URI, HM)

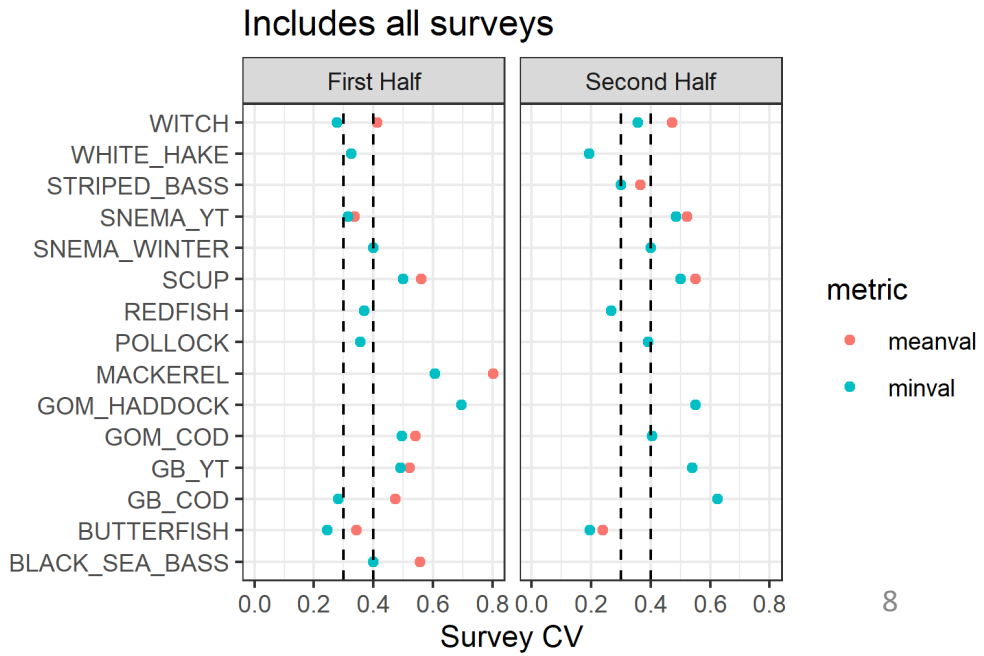
TORs

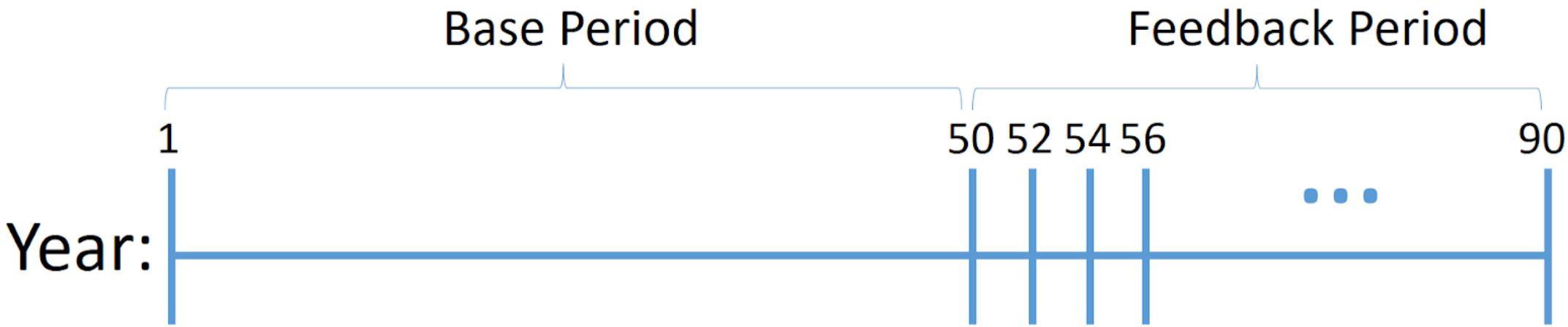
1. Develop methods to create data that if assessed with standard age-based approaches (e.g., VPA or ASAP) could exhibit a strong retrospective pattern. — Make Data
2. Identify a number of index-based methods and a range of harvest control rules for use in closed-loop simulation, using index-based data resulting from ToR 1. — Pick IBMs
3. Identify metrics from the index-based assessment results that could be used in evaluations of trade-offs in performance among harvest control rules and index-based methods. — Select Metrics
4. Evaluate the combinations of index-based methods and control rules using the metrics in ToR 3 to determine candidates for consideration by the Councils or other management authorities. — Crank Sims
5. Provide guidance on specific situations that are and are not well-suited for a particular control rule or index-based method identified in ToR 4. — Advise
6. Create guidelines for setting biological reference points for index-based stocks. — Ref Points

Make Data

- Groundfish-ish
- WHAM closed-loop simulations
- Retros
- Scenarios
- Workflow







Year 1 and Base Period

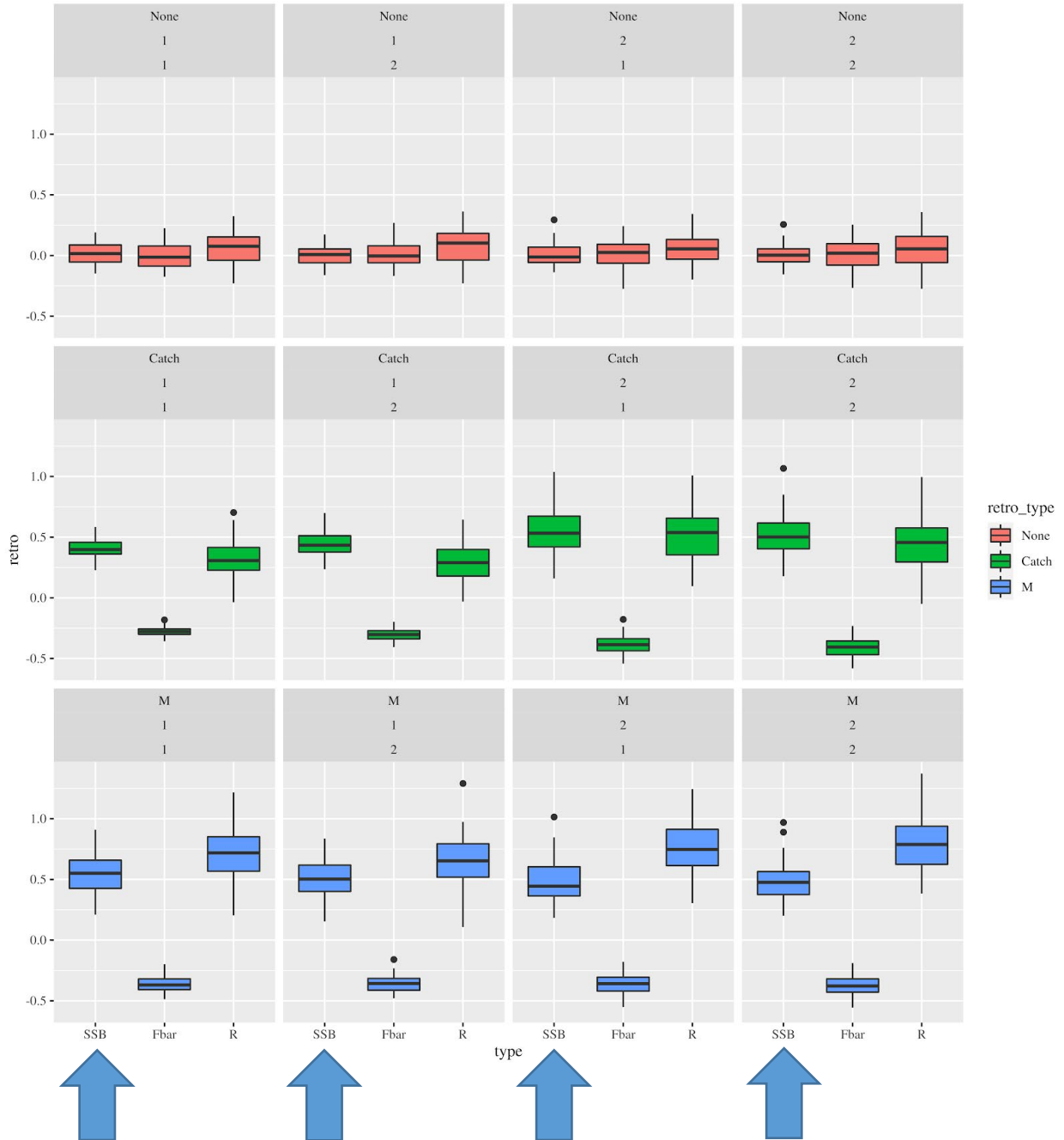
- Run begins
- Equilibrium Pop (Z)
- Fishery develops by scenario
- Recruitment variability only

Feedback Period

- IBM applied every other year
- Catch advice for two years
- Impacts population
- Run ends year 90
  - Collect metrics
  - Repeat 1,000 times

Retros

Distribution of Mohn's rho from 50 simulations



0.0

0.5

0.5

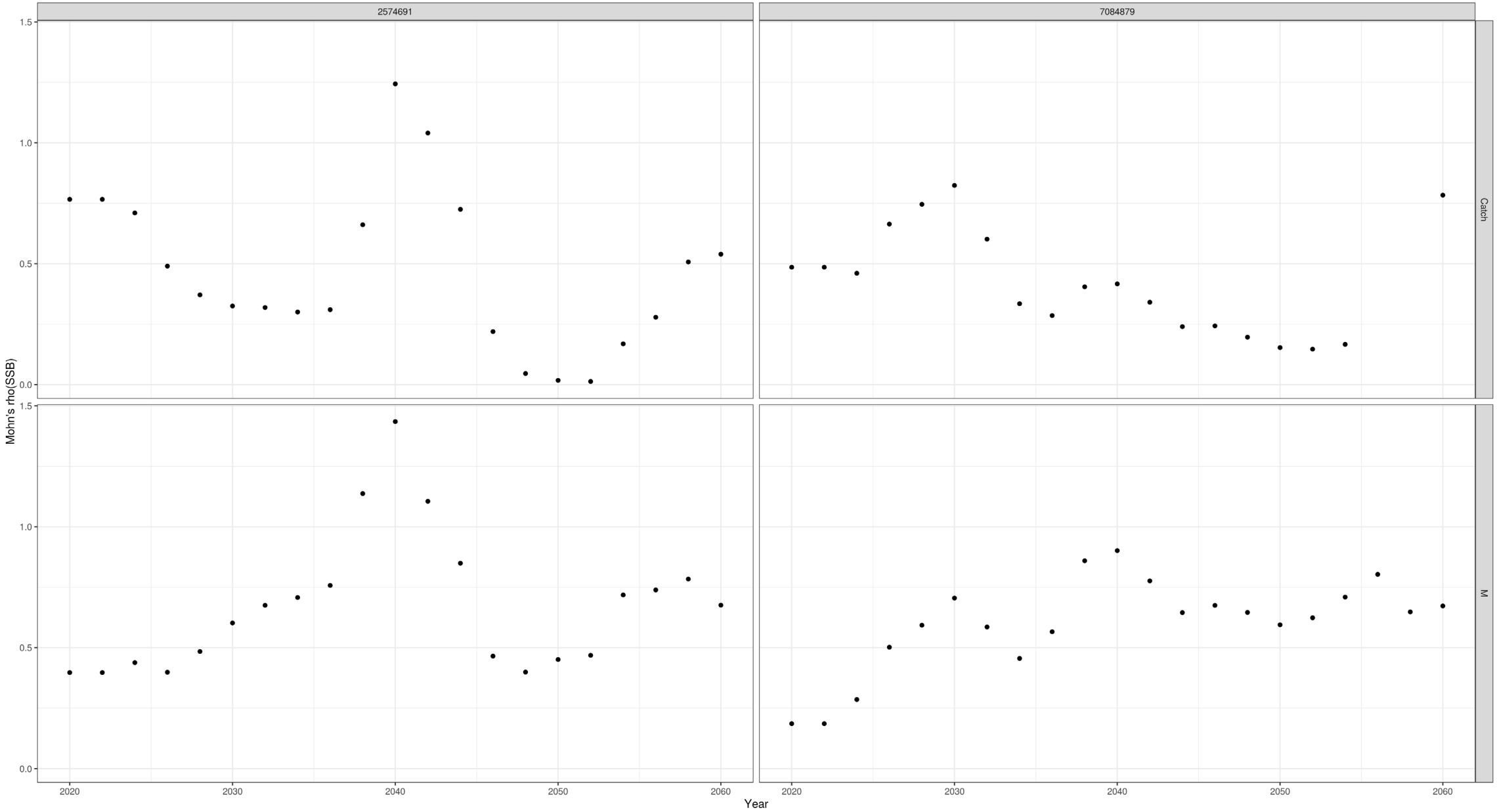
Retro Source

None

Catch
x2.25-5

M
0.32-0.36

Retros in Feedback Period



Retros

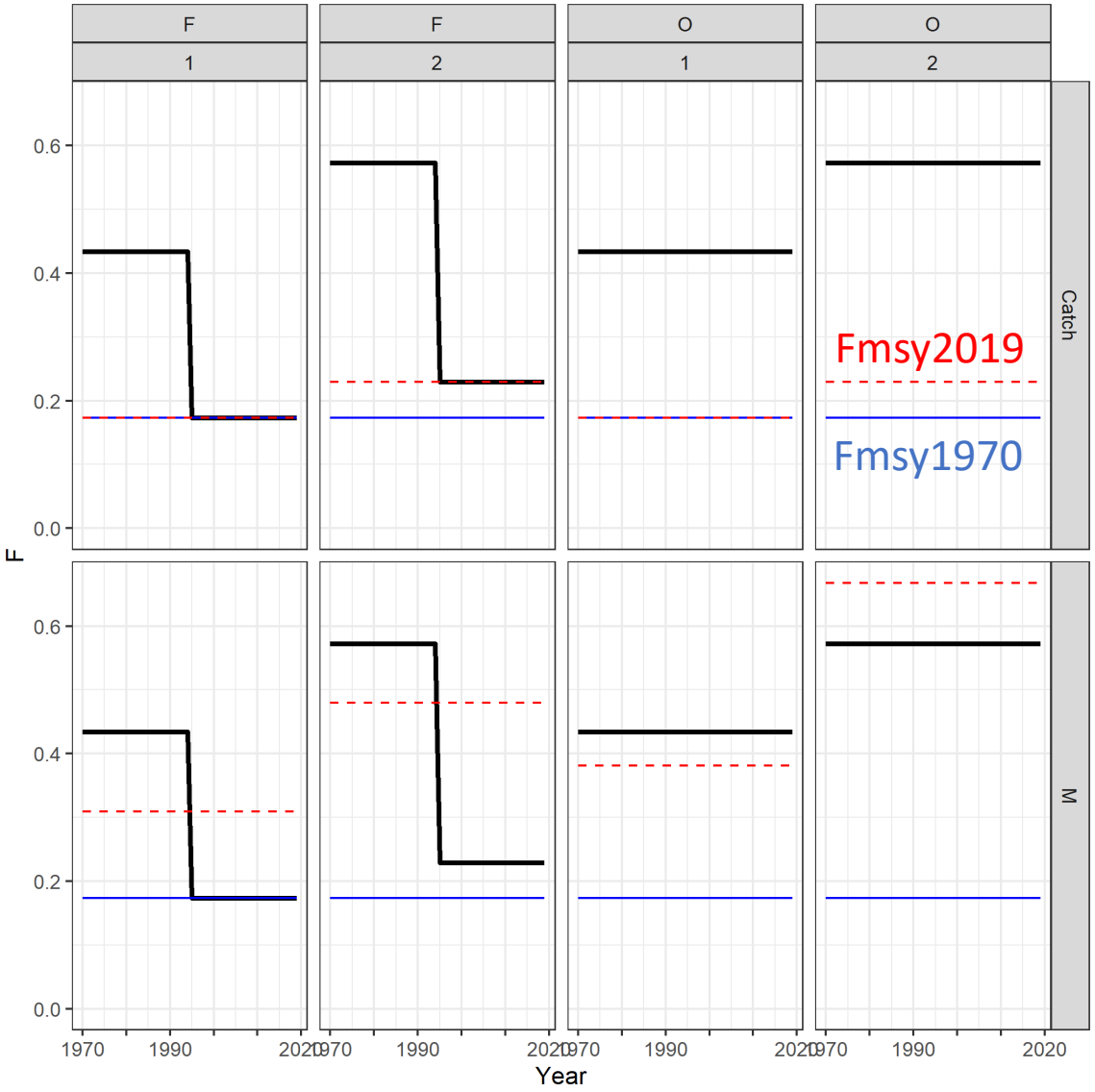
- Why only 2 sources of retro?
  - Catch and M current leading contenders in region
  - Survey q changes did not work well in preliminary explorations
  - Time constraints
- Are these forcing functions realistic?
  - Yes
  - Low current catch for some local stocks means that 5 fold missing could happen (landings, discards, or both)
  - M could more than double (has been estimated to increase 5 fold in some neighboring stocks)
  - But these simulations do not mean that these forcing functions are happening

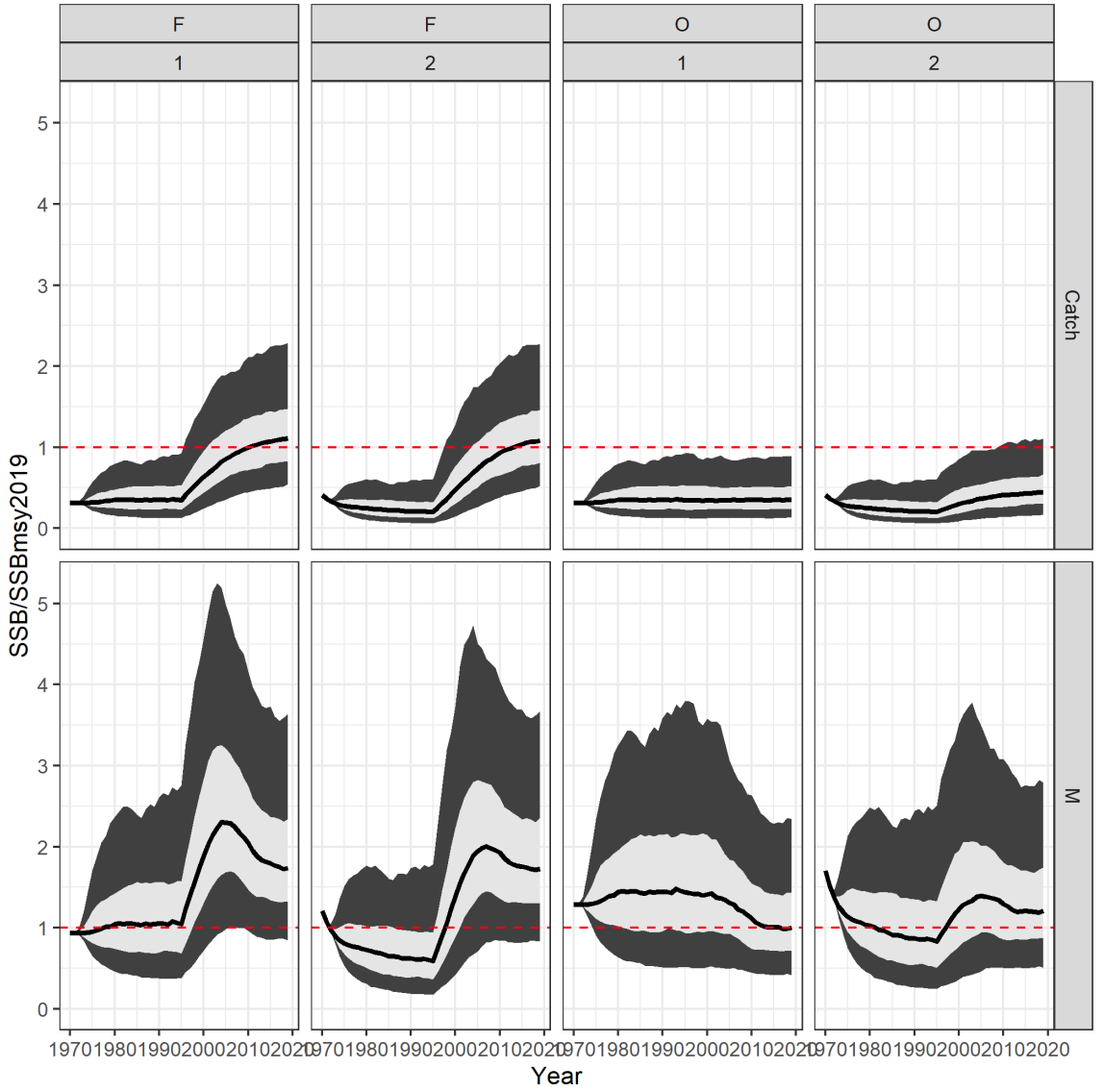
Scenarios

CF1A
CF1R
CF2A
CF2R
CO1A
CO1R
CO2A
CO2R
MF1A
MF1R
MF2A
MF2R
MO1A
MO1R
MO2A
MO2R

Position	Factor	Values
1	Retrospective source	C = catch M = natural mortality N = none
2	Fishing history	F = Fmsy in second half of base period O = overfishing throughout base period
3	Fishery selectivity blocks	1 = constant selectivity 2 = selectivity changes in second half of base period
4	Catch advice multiplier	A = applied as is from IBM R = reduced (multiplied by 0.75) from IBM

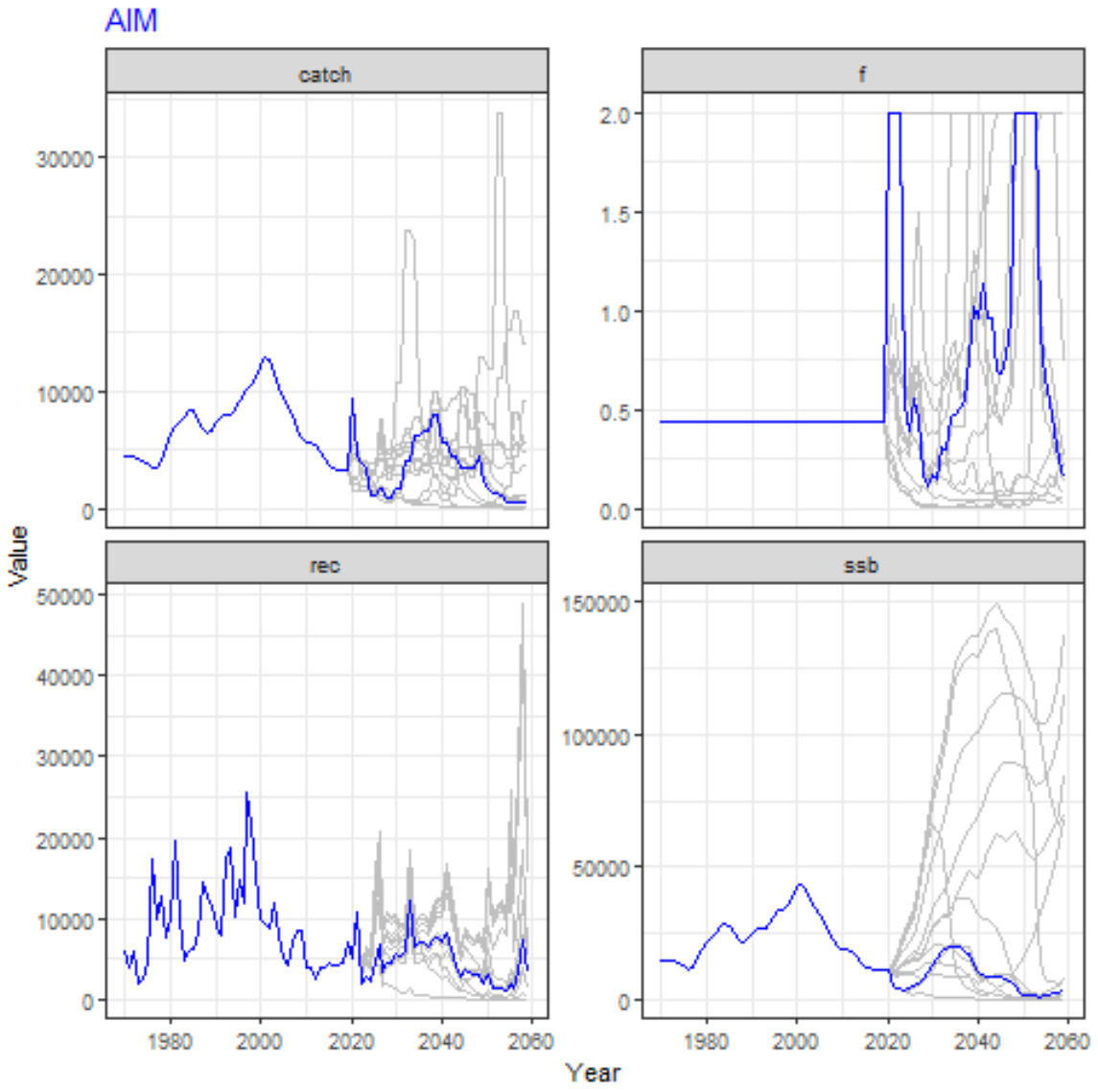
Base Period


F
1
F
2
O
1
O
2
Catch
M
F
0.0
0.2
0.4
0.6
1970
1990
2020
Year
Fmsy2019
Fmsy1970


F
1
F
2
O
1
O
2
Catch
M
SSB/SSBmsy2019
0
1
2
3
4
5
1970
1980
1990
2000
2010
2020
Year

Branching

- Each scenario had 1,000 RNGs
- RNG determined recruitment devs in base and feedback period
- IBMs led to different outcomes based on same scenario and RNG
- Scenario CO1A shown in plot
- OM limited max F to 2.0



Workflow

- GitHub
  - https://github.com/cmlegault/IBMWG
  - Collaborate on code
  - Version control
- Google Drive
  - Store files of results
  - ~300 GB (too big for GitHub)
- Google Docs
  - Collaborate on meeting notes
  - Collaborate on report writing

IBMs

1. AIM = An Index Method *#
2. CC-FM = Catch Curve
3. CC-FSPR = Catch Curve *
4. DLM = Dynamic Linear Model
5. Ensemble method
6. ES-FM = Expanded survey biomass
7. ES-Frecent = Expanded survey biomass *#
8. ES-FSPR = Expanded survey biomass *#
9. ES-Fstable = Expanded survey biomass
10. Islope = common trend based IBM *
11. Itarget = common level based IBM *
12. PlanB (PBS) = survey smoother *#
13. Skate = catch/B driven *#

* indicates member of Ensemble method
# indicates used in regional assessment

FSPR = F40%SPR
FM = F set equal to M
Frecent = average of recent 5 years catch/B
Fstable = F to create stable population

Not Tuned

- All IBMs used in formulaic approach (hands off)
- Real assessments would examine diagnostics from methods
- Future research: dig into results to see if diagnostics would have rejected some catch advice
- Note: DLM < 1,000 simulations due to time constraints

2 Harvest Control Rules

A = applied catch advice directly (treat like ABC)

R = reduced (multiply catch advice by 0.75) (treat like OFL)

- Neither accounts for relative stock size
  - Some IBMs do not estimate relative stock size

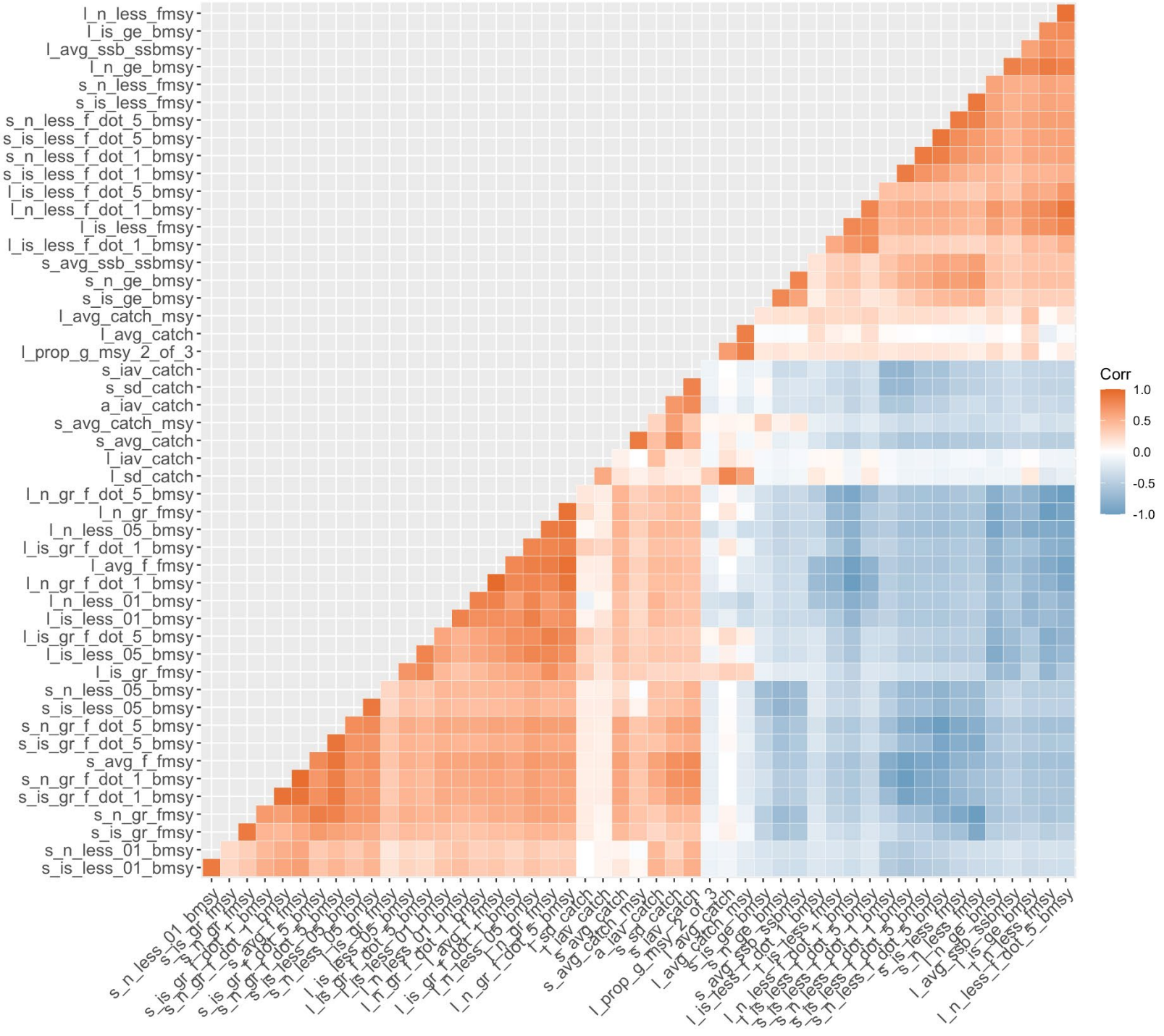
50 Metrics

catch_a_iav_catch
catch_l_avg_catch
catch_l_avg_catch_msy
catch_l_iav_catch
catch_l_prop_g_msy_2_of_3
catch_l_sd_catch
catch_s_avg_catch
catch_s_avg_catch_msy
catch_s_iav_catch
catch_s_sd_catch

f_l_avg_f_fmsy
f_l_is_gr_f_dot_1_bmsy
f_l_is_gr_f_dot_5_bmsy
f_l_is_gr_fmsy
f_l_is_less_f_dot_1_bmsy
f_l_is_less_f_dot_5_bmsy
f_l_is_less_fmsy
f_l_n_gr_f_dot_1_bmsy
f_l_n_gr_f_dot_5_bmsy
f_l_n_gr_fmsy
f_l_n_less_f_dot_1_bmsy
f_l_n_less_f_dot_5_bmsy
f_l_n_less_fmsy
f_s_avg_f_fmsy
f_s_is_gr_f_dot_1_bmsy
f_s_is_gr_f_dot_5_bmsy
f_s_is_gr_fmsy
f_s_is_less_f_dot_1_bmsy
f_s_is_less_f_dot_5_bmsy
f_s_is_less_fmsy
f_s_n_gr_f_dot_1_bmsy
f_s_n_gr_f_dot_5_bmsy
f_s_n_gr_fmsy
f_s_n_less_f_dot_1_bmsy
f_s_n_less_f_dot_5_bmsy
f_s_n_less_fmsy

ssb_l_avg_ssb_ssbmsy
ssb_l_is_ge_bmsy
ssb_l_is_less_01_bmsy
ssb_l_is_less_05_bmsy
ssb_l_n_ge_bmsy
ssb_l_n_less_01_bmsy
ssb_l_n_less_05_bmsy
ssb_s_avg_ssb_ssbmsy
ssb_s_is_ge_bmsy
ssb_s_is_less_01_bmsy
ssb_s_is_less_05_bmsy
ssb_s_n_ge_bmsy
ssb_s_n_less_01_bmsy
ssb_s_n_less_05_bmsy

Correlations



Simulations

IBMlab	CF1A	CF1R	CF2A	CF2R	CO1A	CO1R	CO2A	CO2R	MF1A	MF1R	MF2A	MF2R	MO1A	MO1R	MO2A	MO2R	NF1A	NO1A
AIM	999	1000	1000	1000	999	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	999	999
CC-FM	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
CC-FSPR	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
DLM	734	535	735	535	733	535	735	535	735	535	734	535	733	535	735	535	NA	NA
Ensemble	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
ES-FM	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
ES-Frecent	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
ES-FSPR	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
ES-Fstable	1000	1000	1000	1000	999	1000	999	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	999
Islope	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
Itarget	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
PlanB	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
Skate	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
SCAA	1000	NA	NA	NA	1000	NA	NA	NA	1000	NA	NA	NA	1000	NA	NA	NA	NA	NA

230,147 simulations

Guide to Appendix 6

Figure Type	Report Example	Base	No retro	SCAA
Number of sims	N/A	1	1	1
Scores	4.1	2-14	188-200	269-281
Boxplots	4.9	15-32	201-218	282-299
Trade off (means)	4.15	33-38	219-224	300-305
1,000 points	4.11	39-96	225-260	306-341
Bagplots	4.12	97-125	N/A	342-359
Scenario panel sorted	4.3	126-131	261-266	360-365
Status	4.19	132-133	267-268	366-367
Confetti	4.13	134-187	N/A	N/A
ANOVA plots	N/A	368-445	N/A	N/A
Heatmaps	5.1	446-450	N/A	N/A

Scoring IBMWG Results

Metrics:

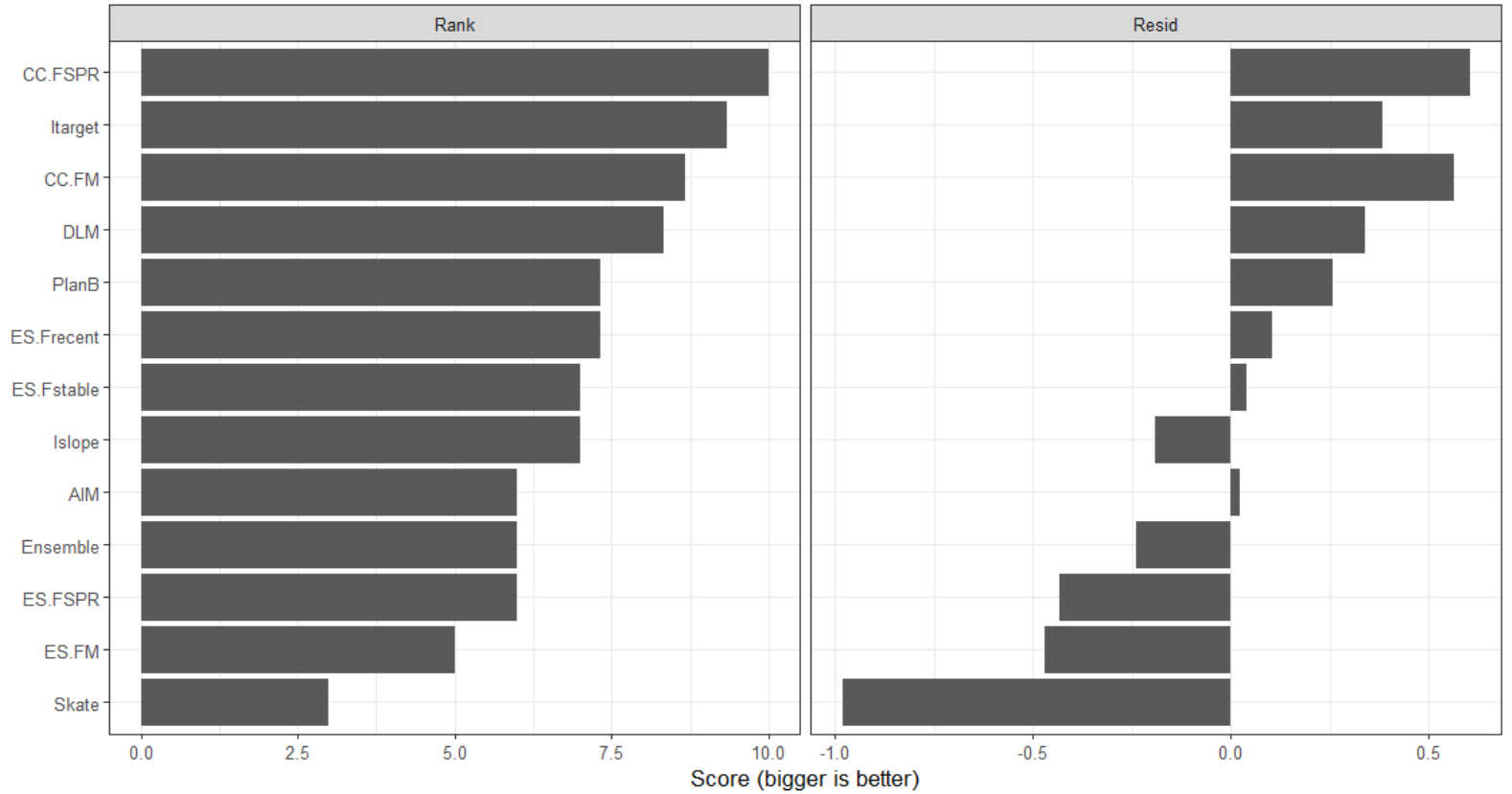
catch_l_avg_catch_msy f_l_avg_f_fmsy ssb_l_avg_ssb_ssbmsy

Set:

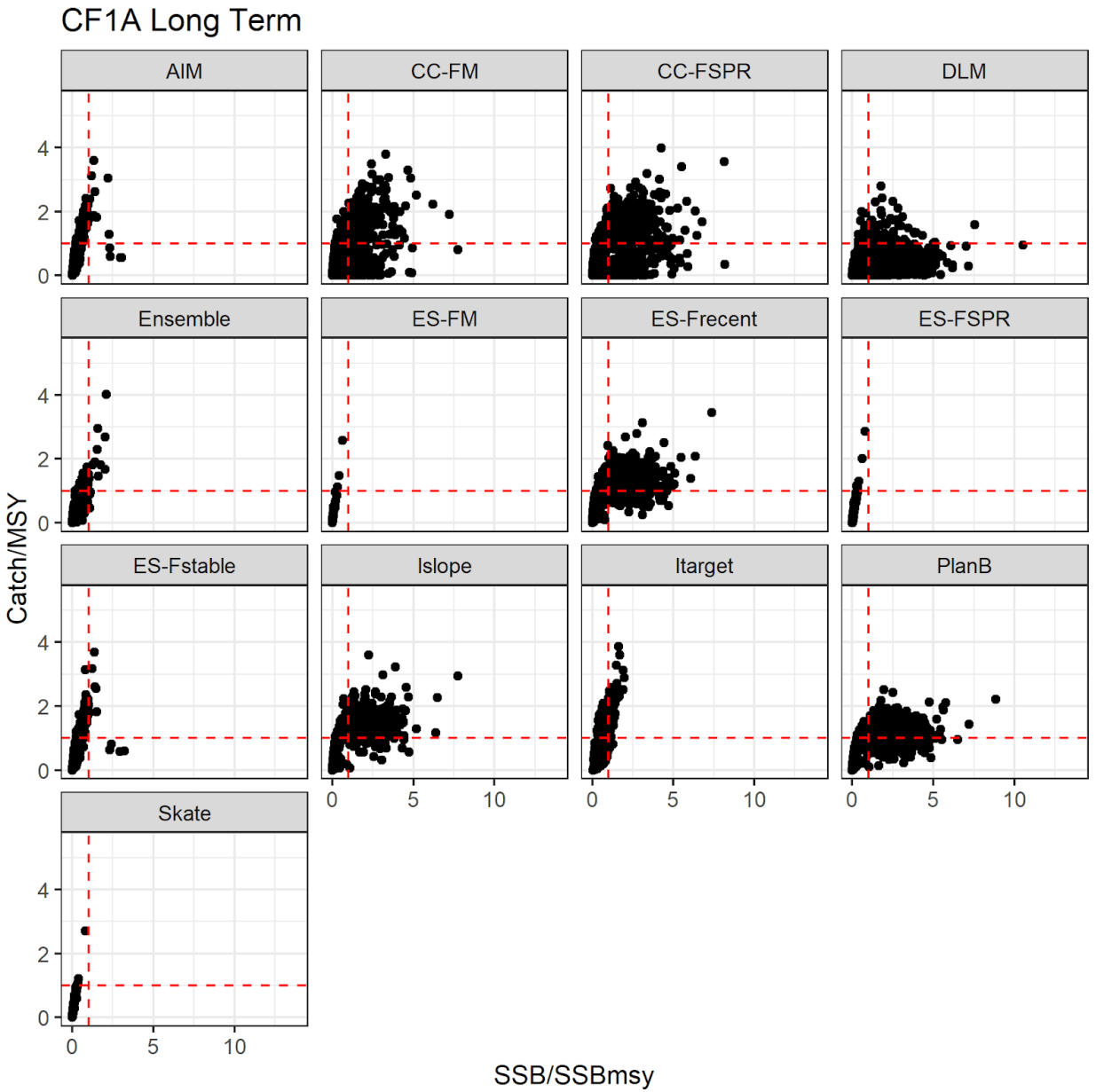
- (•) base
- ( ) noretro
- ( ) scaa

Plot:

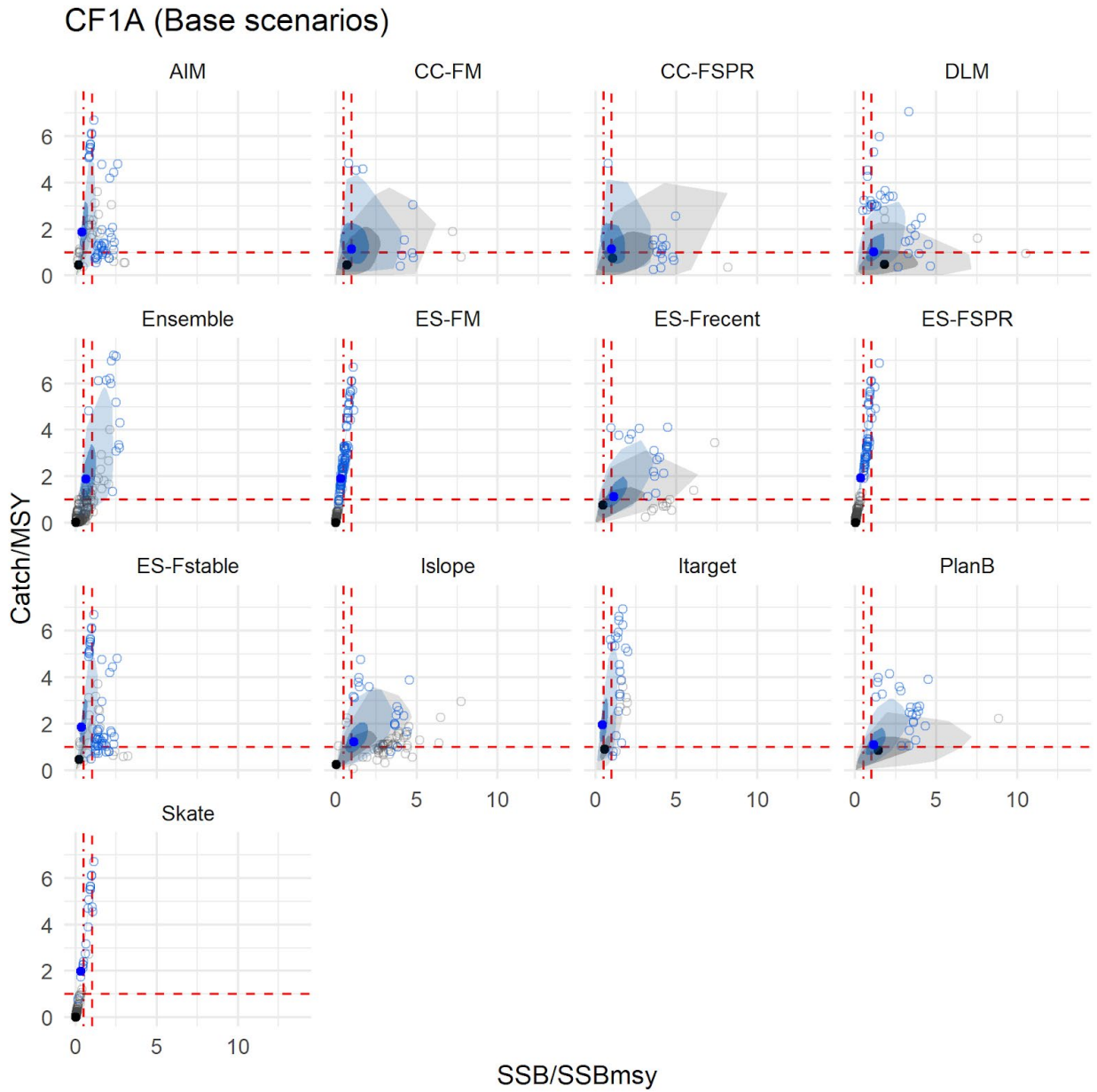
- ( ) Rank
- ( ) Resid
- (•) Both



X	metric	AIM	CC.FM	CC.FSPR	DLM	Ensemble	ES.FM	ES.Frecent	ES.FS
3	catch_l_avg_catch_msy	11	3	5	1	8	9	4	
11	f_l_avg_f_fmsy	5	11	12	13	4	2	9	
37	ssb_l_avg_ssb_ssbmsy	2	12	13	11	6	4	9	



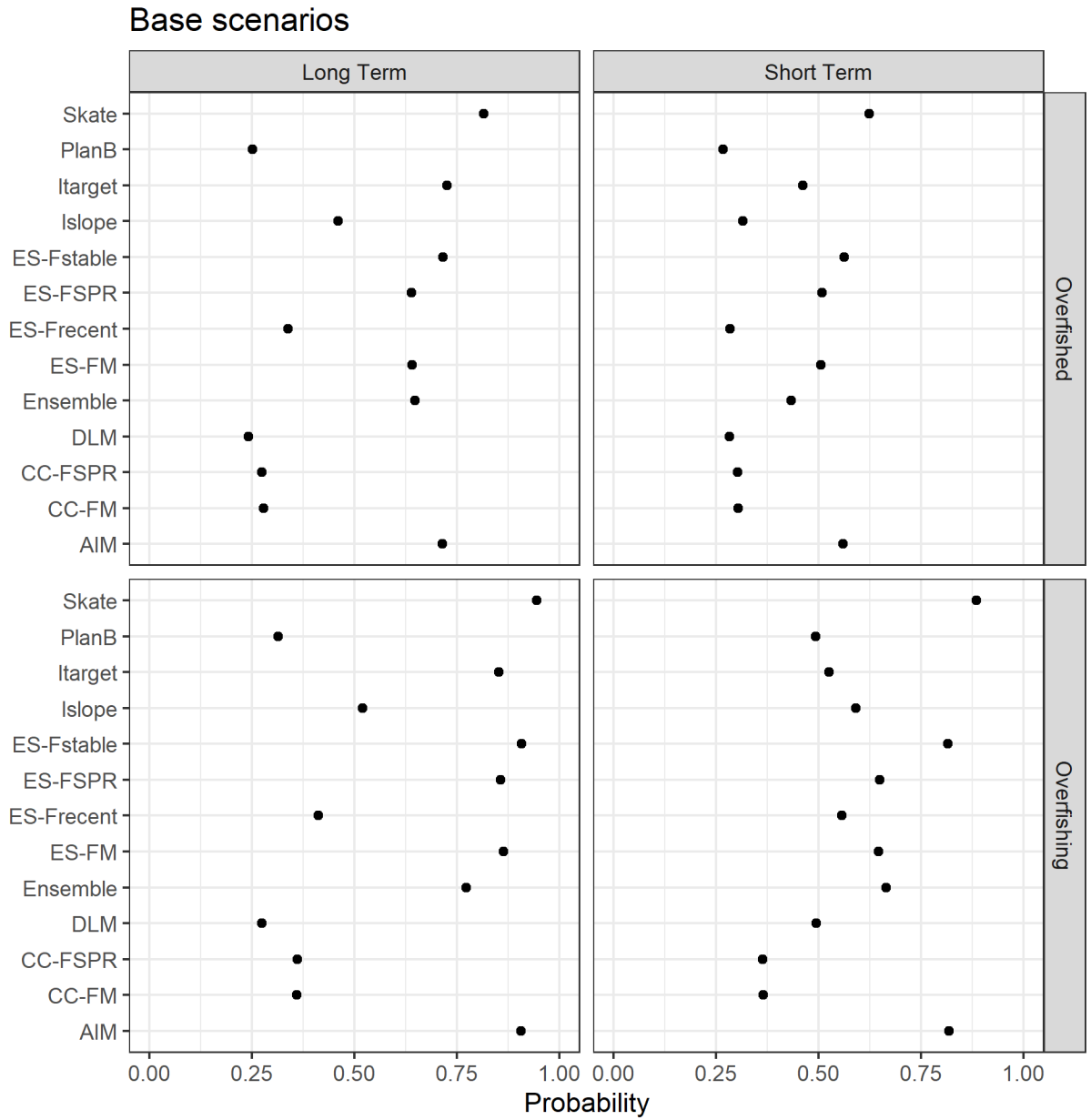
Long term (black) and short term (blue)

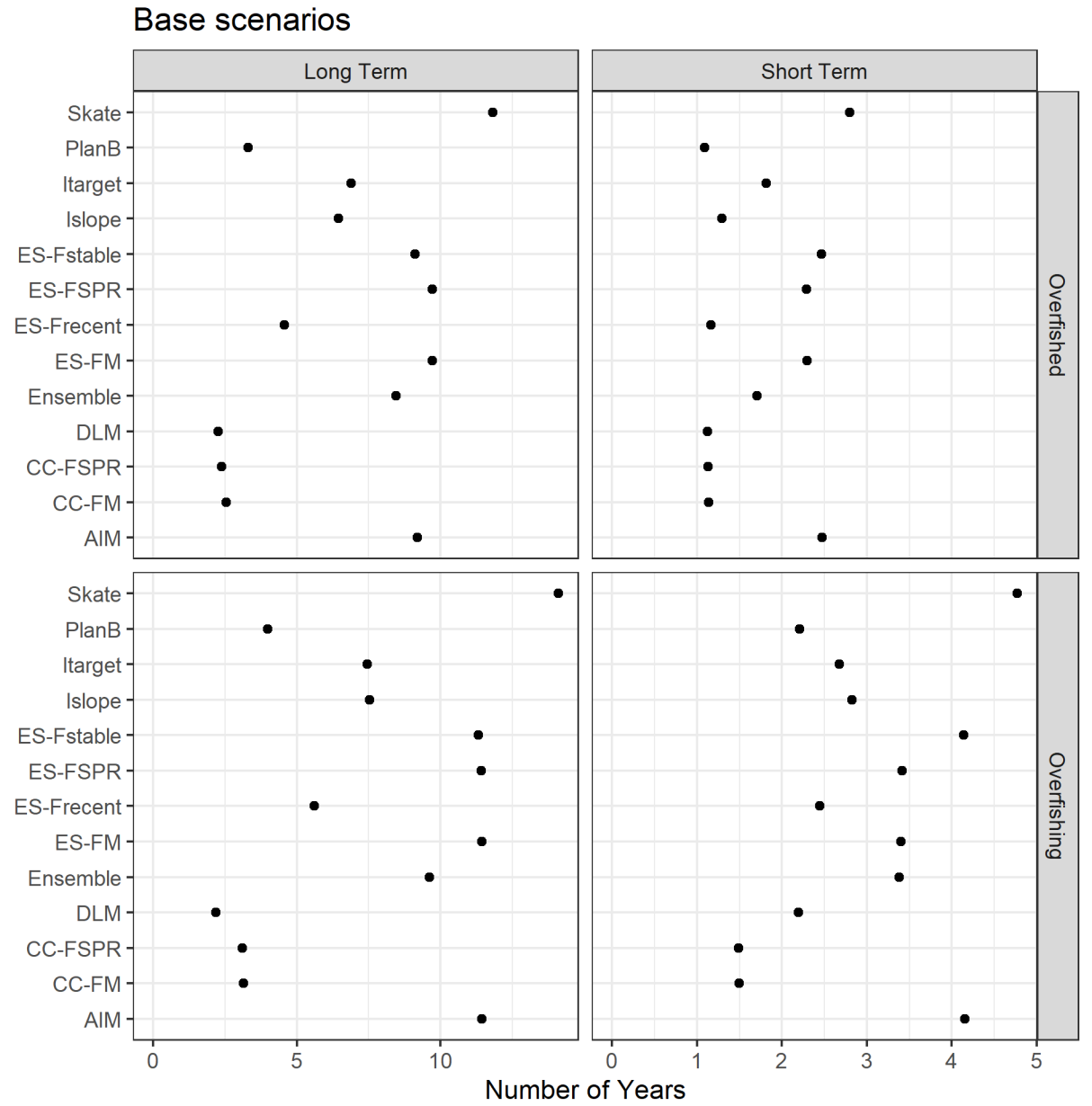


Figs A6.132, A6.133

Risk

Overfished: SSB < 0.5 SSBmsy
Overfishing: F > Fmsy





No Retro

- None retro type closer to 1.0 than either catch or M retro types

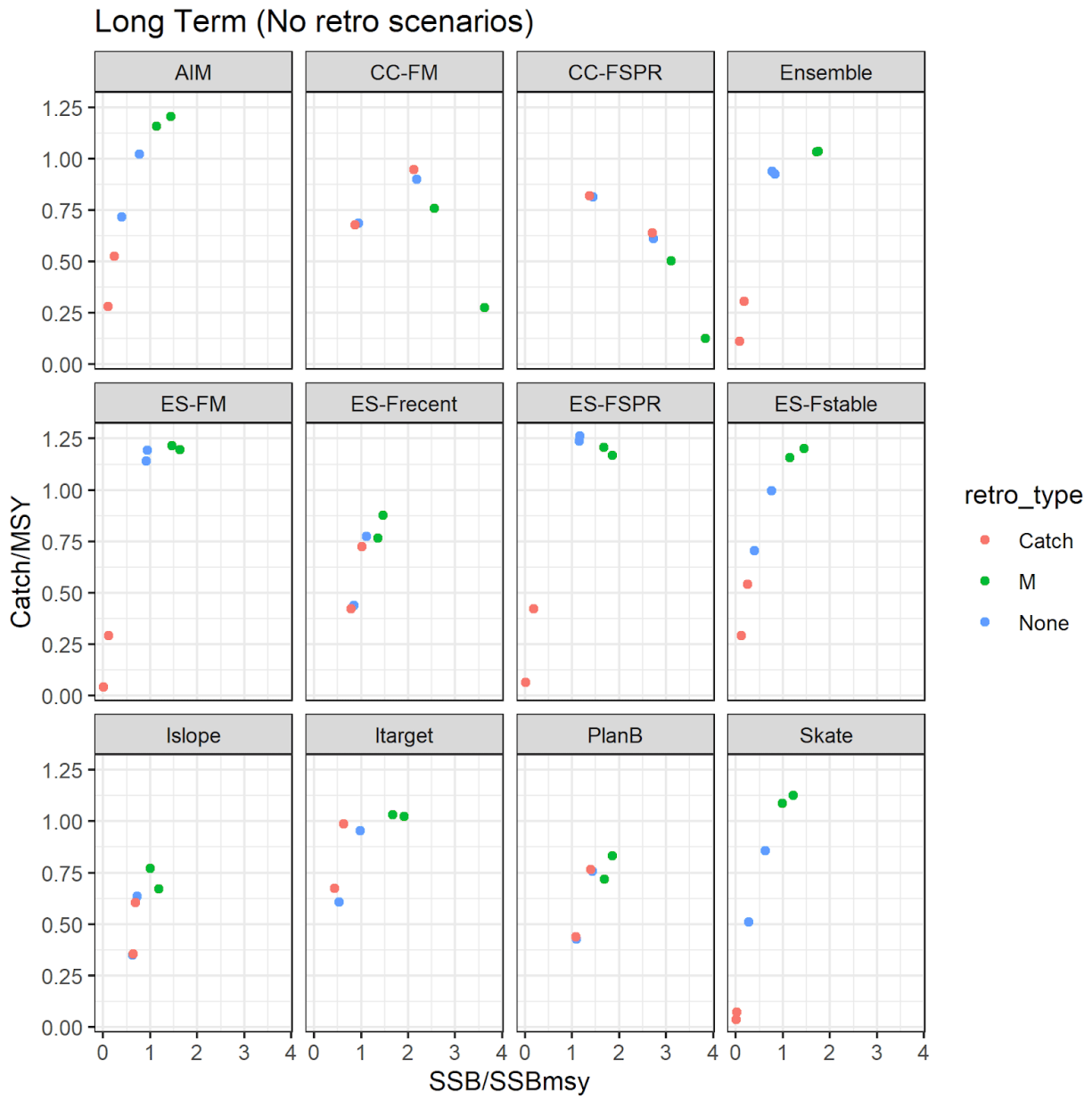
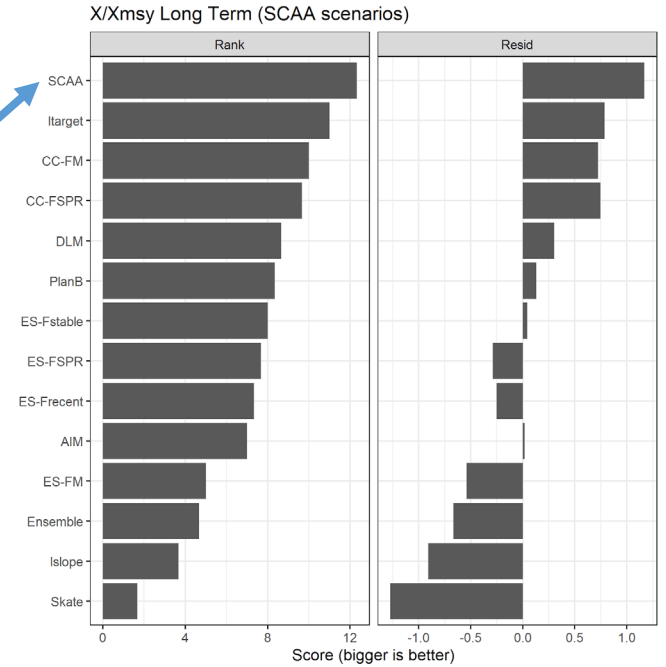
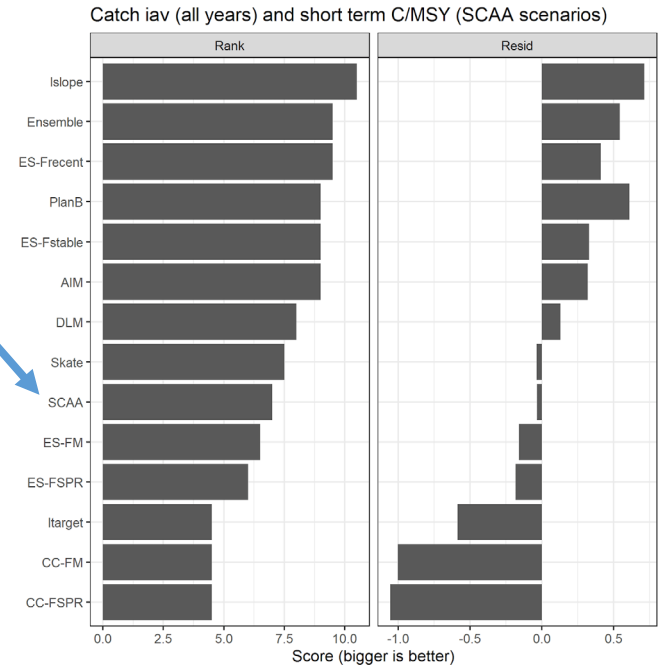


Fig 4.15

SCAA

- SCAA performed well compared to IBMs for many metrics
- There were some metrics where SCAA performed in the middle of IBMs





Figs 4.16, 4.17

SCAA

- Performed well in terms of status

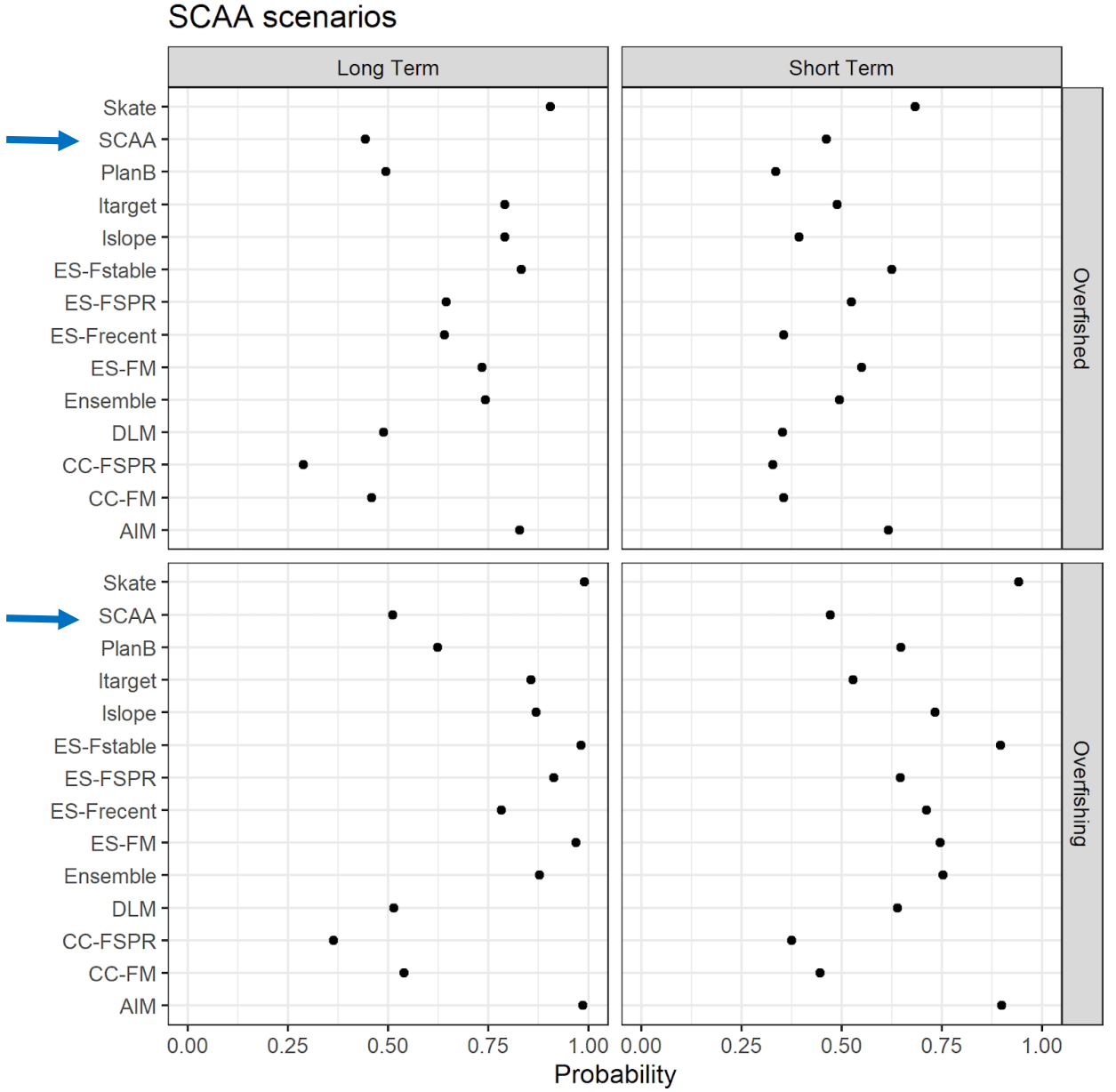


Fig 4.19

Main Results

- Base analyses found two groups
  - CC-FSPR, CC-FM, DLM, PlanB, ES-Frecent, Islope: SSB and F ratios better
  - Skate, AIM, ES-Fstable, ES-FSPR, ES-FM, Ensemble, Itarget: Catch ratios better
- No retrospective source causes long term SSB and F to be closer to MSY values than either the catch or M retrospective sources
  - Alternative to bigger is better approach to metrics should be considered in the future
- Overall, none of the IBMs outperformed SCAA with rho-adjustment
- When an SCAA is rejected due to retro, should not expect IBM to perform better than rejected model
  - Also lose the use of additional data, status determinations, and hypothesis testing with IBMs

Other SCAA rho-adjustment results

- Brooks and Legault 2016 https://doi.org/10.1139/cjfas-2015-0163
- Wiedenmann and Jensen 2019 https://doi.org/10.1139/cjfas-2018-0129
- Legault 2020 https://doi.org/10.1093/icesjms/fsaa184
- Rho-adjustment not perfect
- Replacing SCAA exhibiting strong retro with IBM does not produce improved catch advice

Caveats

Caveat	Address in future using this framework?
• Groundfish-ish	• Yes
• Single source and magnitude of retro	• Yes
• Changing forcing function for retro over time	• Yes, with mods
• Assessed every other year	• Yes
• Hands off applications of IBMs	• No
• Limited methods to derive catch advice	• Yes, with mods
• Limited formulations of IBMs	• Yes, with mods

Ref Points Challenge

- IBMs typically do not have production functions that allow evaluation of trade offs between catch and future population size
  - Standard approach to deriving reference points in age-based models
- Instead IBMs may attempt to find thresholds in the observed survey and catch time series that can serve as reference points
  - These should not be thought of as MSY proxies
- Automating a search for such thresholds is challenging
- This work provides a trove of information that could be mined for such thresholds
  - Time constraints prevented this working group from doing so

Recommendation

- “As with all research, many questions were raised during this study. The framework developed for these simulations is well suited to address many of them immediately or with minor modifications. The IBMWG recommends this work be continued to explore the results generated during this study as well as building on these results to address new questions.”
  - From Executive Summary

Lessons Learned

- Need to have sufficient time for large simulation studies
  - Time to think is an important part
  - TORs need to match available time
- Collaboration during pandemics is possible and fun
  - Many tools facilitated this work
  - Treasure trove of information still to be explored
- Current approach of rejecting age-based models with strong retros and replacing with index-based may not be producing desired results
- Framework developed can be used to explore specific situations

Questions?

chris.legault@noaa.gov

Data Portal:

https://apps-nefsc.fisheries.noaa.gov/saw/sasi/sasi_report_options.php

Assessment Year: 2020

Species Name: Index Based Methods

Stock Area: Georges Bank

Review Type: Research Track