Stakeholder Workshop on Management Reference Points for Gulf Menhaden Fisheries

Hosted by

Gulf States Marine Fisheries Commission

Facilitated by

Dr. Michael Jones

July 17-19, 2019

Hilton New Orleans Riverside
New Orleans, LA

July 2019 Pub No. 292

COMMISSIONERS AND PROXIES

ALABAMA
Chris Blankenship, Commissioner
Alabama Department of Conservation
and Natural Resources
64 North Union Street
Montgomery, AL 36130-1901

Proxy:
Scott Bannon, Director
ADCNR, Marine Resources Division
P.O. Box 189
Dauphin Island, AL 36528

Representative Steve McMillan
P.O. Box 337
Bay Minette, AL 36507

Chris Nelson
Bon Secour Fisheries, Inc.
P.O. Box 60
Bon Secour, AL 36511

FLORIDA
Eric Sutton
FL Fish and Wildlife Conservation Commission
620 South Meridian Street
Tallahassee, FL 32399-1600

Proxy:
Dan Ellinor
FL Fish and Wildlife Conservation Commission
620 South Meridian Box 4B2
Tallahassee, FL 32399-1600

Representative Jay Trumbull
317 House Office Building
402 South Monroe Street
Tallahassee, FL 32399-1100

LOUISIANA
Jack Montoucet, Secretary
Louisiana Department of Wildlife and Fisheries
P.O. Box 98000
Baton Rouge, LA 70898-9000

Proxy:
Jason Froeba
Louisiana Dept. of Wildlife and Fisheries
P.O. Box 98000
Baton Rouge, LA 70898-9000

Senator R.L. "Bret" Allain, II
600 Main Street, Suite 1
Franklin, LA 70538

John Roussel
1221 Plains Port Hudson Road
Zachary, LA 70791

MISSISSIPPI
Joe Spraggins, Executive Director
Mississippi Department of Marine Resources
1141 Bayview Avenue
Biloxi, MS 39530

Read Hendon
USM/Gulf Coast Research Lab
703 East Beach Drive
Ocean Springs, MS 39564

TEXAS
Carter Smith, Executive Director
Texas Parks and Wildlife Department
4200 Smith School Road
Austin, TX 78744

Proxy:
Robin Riechers/Lance Robinson
Texas Parks and Wildlife Department
4200 Smith School Road
Austin, TX 78744

Troy Bello Williamson, II
P.O. 967
Corpus Christi, TX 78403

TABLE OF CONTENTS

The two-day workshop is available on the GSMFC YouTube channel at:

WORKSHOP VIDEOS

- DAY 1 – 1:00 TO 4:30
- DAY 2 – 8:00 TO 5:00
- DAY 3 – 8:00 TO NOON

STAKEHOLDER WORKSHOP ON MANAGEMENT REFERENCE POINTS FOR GULF MENHADEN FISHERIES

July 17 - 19, 2019
New Orleans, LA

Facilitator
Mike Jones - Michigan State University - East Lansing, MI

Participants
Doug Butterworth – Univ. of Cape Town – South Africa
Jason Adriance - LDWF - New Orleans, LA
Amy Schueller – NMFS – Beaufort, NC
Robert Leaf - USM - Ocean Springs, MS
John Mareska - AMDR - Dauphin Island, AL
Jerry Mambretti - TPWD - Dickinson, TX
Trevor Moncrief - MDMR - Biloxi, MS
Peter Himchak - Omega Protein - Tuckerton, NJ
Ben Landry – Omega Protein - Houston, TX
Francois Kuttel - Westbank Fishing - New Orleans, LA
Chad Hanson - Pew Charitable Trust - Crawfordville, FL
Marianne Cufone – Recirculating Farms – New Orleans, LA
Ed Swindell – Marine Process Services - Hammond, LA
Scott Herbert – Daybrook Fisheries – New Orleans, LA
Randy Pausina – LA CCA - Covington, LA
Camp Matens – LA CCA - Baton Rouge, LA
Shaun Gehan – Omega Protein – Washington, D.C.
Kendall Dix – Healthy Gulf - New Orleans, LA
Charlotte Runzel – National Audubon Society – Washington, DC
Alyssa Conti – Green Justice, Loyola Law, New Orleans, LA
Jennifer Stucker – Green Justice, Loyola Law, New Orleans, LA

Staff
David M Donaldson - GSMFC - Ocean Springs, MS
Steve VanderKooy - GSMFC - Ocean Springs, MS
Debbie McIntyre - GSMFC - Ocean Springs, MS
Joe Ferrer - GSMFC - Ocean Springs, MS

15:30

Welcome and Introductions

Dr. Mike Jones served as the facilitator in the February Reference Points Workshop (*RP Workshop henceforth*) and provided some background. **Jones** is a professor at Michigan State University and codirects the Quantitative Fisheries Center. **Jones** has done work in the Great Lakes as well as the Pacific Northwest including Alaska. Four years ago, **Jones** facilitated the ecosystem management goals workshop for Atlantic Menhaden on behalf of the ASMFC. The meeting will break today around 4:15

pm so that the analysts can review a couple of last minute items that they worked on for the results portion of this meeting. Participants are welcome to stay but the discussion will be short and technical in nature.

Everyone attending is considered a participant in this process and is encouraged to ask questions and contribute to the discussion throughout the two days. There may be folks attending via the GoToMeeting to allow Public Comment and so that the online audience can ask any questions as well. **Jones** did remind everyone to be respectful of everyone's opinions and be hard on the problem, not the presenter.

The participants introduced themselves around the table and room and provided a little of their background for their inclusion in this process.

Agenda Overview

The agenda was provided which had been updated slightly and sent out a week ago. Most of this afternoon was background for how we got here. **Jones** introduced Dr. **Butterworth** who will summarize the work that went on prior to this meeting. He went over the agenda and explained the overview and workshop goals and summarized where the collective group ended up after the RP Workshop. Throughout we will be discussing a practical Management Procedure (MP).

Hanson noted that the original reason to meet and get started on reference points was to address the recommendations that came out of the last stock assessment review. We discussed extensively at the RP Workshop, potential reference points that could be used in the assessment to address stock status but this workshop and the agenda appears to <u>not</u> be an exploration of RPs associated with assessed biomass, but rather only those associated with a biomass index. There were a number of other reference point options discussed at the RP Workshop and he thought that they would be 'tested' as well to see how they performed using the MSE process. How is THIS effort going to be used in the stock assessment? There needs to be a legitimate way to gauge the status of the stocks and Hanson felt this was the primary reason for working on reference points, not to set catch limits only.

Jones thinks that the key question raised by this comment is whether the reference points in the MP are based on an adequate measure of the status of the stocks. This should become part of the discussion about how the harvest control rules were developed, and participants should make sure the information is made clear during those upcoming discussions.

Jones reviewed the RP Workshop outcomes. The intent for that meeting was: to discuss the purposes for reference points; to identify candidate objectives for the fishery; to review current Menhaden status; to assess candidate reference points. He summarized the fundamental objectives agreed upon at the RP Workshop: balancing fishery and ecosystem needs, making sure user groups accept responsibility, and that all have confidence in management. The need to address reference points was one of the concerns raised by the reviewers of the last SEDAR benchmark assessment. The RP Workshop put that into perspective and generated a pathway to get to something to inform management decisions. Another motivation was the desire by the industry to have reference points adopted and work toward harvest control rules to satisfy the Marine Stewardship Council (MSC) certification process.

Jones showed the list of 'fundamental' and 'means' objectives that came out of the RP Workshop along with the conclusions generated by the participants (below). At the end of that workshop, there was a plan for the analysts to team up to move forward with the empirical index-based reference points using the MSE approach to evaluate the forecast performance of the alternative references points suggested during the RP Workshop. **Butterworth** and **Rademeyer** had provided an initial 'proof-of-concept' for an empirical reference point for the RP Workshop. Since that time, they have further revised it after evaluating some of the other options. It was further agreed that the group would consider other possible reference points in the future when ecosystem modeling results are more informative.

Objectives: Fundamental

- Balance needs of fishery and needs of ecosystem to maintain long-term sustainability
- All user groups accept shared responsibility for maintaining and improving ecosystem health, population abundance, and biodiversity
- All partners and general public have confidence in the sustainability of the fishery and the industry, and in management

Objectives: means

- Adequate SSB to ensure high recruitment
- Minimize negative effects on predators and habitat
- Minimize bycatch
- Maintain sustainable commercial fishery
- Use stock assessment data to ensure maintenance of sustainable stocks
- Allow management flexibility
- Take environmental factors into consideration
- Maintain/restore historic range, age structure and productivity
- Improve monitoring and assessment procedures
- Have management regime in place sufficient to achieve other objectives

Hanson noted that **Dr. David Chagaris** (Univ. of FL) was not present and asked if he would have any update on his ecosystem work for the MAC meeting in the fall. Hanson was under the impression that **Chagaris** was close to having something ready to present. **VanderKooy** stated that **Chagaris** (who is one of the members of the analyst team but was unable to attend this workshop) may have a presentation on his EwE modeling efforts ready for the fall MAC meeting along with a couple others in the ecosystem modeling world (Dr. Skyler Sagarese, Matt Nuttall, and Dr. Kim de Mutsert). **Schueller** reported that Chagaris was integral in work with ecosystem approaches in Atlantic Menhaden. She stated on the Atlantic side that they are very near ready to present a suite of recommendations for consideration on the Atlantic and will go through a CIE review. The ecosystem based approach for Menhaden along the Atlantic has been a very long process. **Chagaris** is working on something similar for the Gulf but the Gulf work is not near as far along and it has not been evaluated yet nor is it really ready yet. **Jones** admitted that developing approaches like this can take a really long time so the expectation that this process could be done quickly for the Gulf is not realistic. This has been ongoing from 2012 to 2019 and is not something this group should expect soon. **Hanson** noted that a lot of meetings end with suggestions but nothing happens after. On the Atlantic, they have a technical group to follow-up on these items and the Gulf might want to form a working group to work on ecosystem modeling and review the progress made on these types of modeling.

Himchak mentioned that in the discussions at the RP Workshop, it was generally agreed that the concept of the index-based reference points was better than the other options so that is the approach that was taken. This was the only approach left standing at the end of that meeting. **Hanson** suggested that there should be other alternatives as well. We also looked at SSB related alternatives that were more correlated with the stock assessment results. **Hanson** had the expectation that those would also be addressed in this meeting through the MSE process. **Jones** suggested that he should revisit this should it not be answered satisfactorily by the end.

Workshop Goal

To review and discuss the results of a management strategy evaluation for Gulf Menhaden.

Background

Butterworth showed a video "On the Benefits of Harvest Strategies" Butterworth MSE Video explaining the MSE (Management Strategy Evaluation) process and how it is applied in marine fisheries around the world. **Butterworth** noted that the conventional assessment approach to coming up with catch limits versus the MPs approach are not contradictory. MPs do not ignore assessments, rather, assessments are the basis for the tests of the MPs. As a bit of history, the need for MPs came up because of the uncertainty of the 1970s and 1980s about sustainability of whaling. They had established harvest control rules, but a decade later determined that they were not really working as they had hoped so the MP approach was developed. The original decisions had been based upon the assessment and the 'best data available' but it was not good enough. What they did instead was look at the consequences of the harvest control rules under a wide variety of conditions. They ultimately agreed that it was better to go with the most extreme cases to minimize risk for the population so that even if the assessment was inaccurate, they would err on the side of caution. The MP approach takes what is beyond the typical assessment because takes uncertainties into account, but it does not desert the assessment.

Jones tried to connect the dots between **Butterworth's** presentation and **Hanson's** concerns about the other reference points that had been discussed at the RP Workshop. **Hanson** had thought we would be considering both the survey indices and estimates of the SSB as candidates for RP. **Hanson** wanted to know the correlation between the indices and the assessed biomass. Certain indices may be misleading due to their uncertainty, particularly if they are biased. **Butterworth** noted that it is common to use the information that has the most impact on the assessment in your MP, especially for an empirical control rule. The reality is that the indices of abundance drive the assessment, so being able to use the same indices in the MP adds confidence. A MP is like an autopilot but not where you do not have a driver, it is like an airplane where the pilot watches and can make adjustments. The assessment needs to be updated on a regular basis, more like a 'turn-the-crank' to regenerate the indices and be sure that we are not getting out of sync with the population trends. The intent is not to conduct a benchmark frequently, but rather have annual diagnostics to review the indices and determine if the MP is performing as it should. Again, the assessment is what you use as the foundation for the simulation testing.

Jones noted that it would be worth looking at how the indices correlate since they tend to drive the assessment as **Hanson** pointed out earlier. **Schueller** will dig out that information from the benchmark for the group. In the meantime, **Butterworth** will summarize the MSE process which was provided going into the RP Workshop.

Description of MSE Process

Butterworth reviewed the *January Paper* he and **Rademeyer** had put together for the RP Workshop. He noted that a principal reason for doing this is to support the industry's desire to satisfy their MSC Certification. Part of the Certification is the requirement for some form of harvest control rule (HCR) that forces you to agree to the management rules before you play the game. Having an HCR in place triggers an immediate response and helps to ameliorate any potential problems in a population.

A second reason is to begin to develop an HCR that includes ecosystem considerations. Given the current level of understanding, this will not resolve the issue forever, but is instead a work in progress which can incorporate new insights as they emerge in the future. **Butterworth** provided an example from the *January Paper* (Figure 1 here). There is no evidence that there have been ecosystem issues associated with Gulf Menhaden abundance since 1992, so one strategy would be to suggest that we at least do not want biomass to go as low as the minimum value during that period in 1992. Starting in 2017, they projected how the SSB index would be expected look on average for the next 20 years (Figure 1) if there was no HCR (red curve) and what would happen with a candidate HCR in place (black curve). As we proceed, we will work through the uncertainties as they were tested. (See the "*January Paper*" for further details).

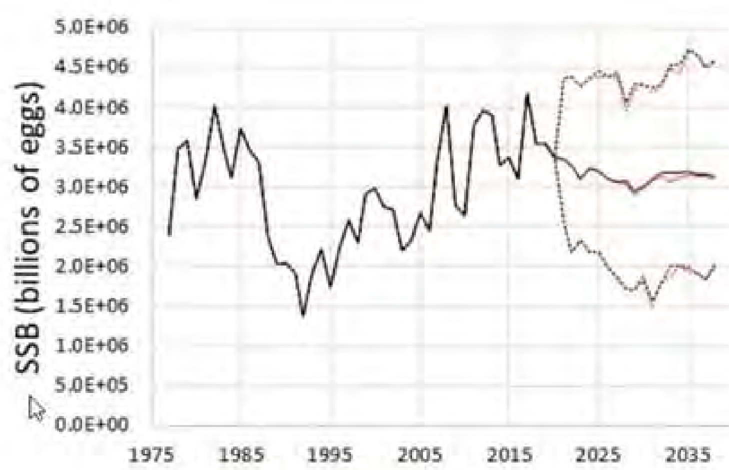


Figure 1. Historical estimates and projected 20-year median and 90%iles for a series of quantities for the **robustness test** with **10 years bad recruitment** with **Fmax=4.0**, without (red lines) and with (black lines) the management rule (Figure 4 from the *January Paper*).

Dix asked whether we have data prior to 1975 for Menhaden. **Schueller** answered that we do but not for index data. The gill nets did not include Menhaden in the catches consistently until the 1980s but the other fishery-independent data became much more consistent in the mid-1970s, therefore we began the assessment at 1977 which is where these graphs start.

Hanson asked if we have any idea what caused the decline reflected on the chart around 1992 in the graph above. Was it an environmental issue, fishing effort, etc? What is ultimately driving recruitment? **Schueller** pointed out there has been some work regarding the environmental pieces but it is not conclusive at all. The assessment team included Mississippi and Atchafalaya River discharges (that were proposed at one time as strong indicators) as sensitivities in the assessment but did not end up carrying much weight in the end. **Leaf** explained that there are several papers out regarding environmental indicators but honestly, the data comparing to the 1950s and 1960s is sparse and just does not show much at all. Historic environmental information that explains variation in Menhaden abundance is just not there. **Schueller** stated that there just is no silver bullet that we can point to for the reduced SSB at that time, it is likely a combination of factors but to try to say something definitive would be more arm waving than anything. **Butterworth** said that regarding the low point in SSB, there was a period of relatively low recruitment but only during that particular time period. The last decade or more has been much higher. One variable which does tend to work on recruitment is temperature. A stock declines when it is near the high end or low end of its preferred temperature because it is struggling because it does not like the temperature. In the middle of a temp range, there are a number of factors that can come into play. **Dix** tried to summarize this discussion: over roughly a seven-year period we lost around two-thirds of the stock spawning biomass and we do not know why. **Butterworth** indicated that abundance was low because recruitment was low but we do not know if it is a cause or an effect and that is probably all we can say. **Jones** noted that despite it being low at that point, it did increase immediately following so it was not a continual decline.

Cufone said, in looking forward, if we are not sure if its environment or landings, are we taking into account a potential increase of landings as aquaculture develops and MSC develops in what we expect to see if future recruitment? **Butterworth** stated yes, we are coming up with a proposal that says the MP will put restrictions on the fishery should things get bad. If you start seeing catches going outside the range (exceptional circumstances), there will be an action taken to resolve the issue. **Moncrief** stated as a follow-up for this conversation that the time period then was one of fluctuation in the number of plants and vessels operating and increased tropical activity, especially in 1992 with Hurricane Andrew. The take home is that we <u>do not</u> want to get back to that point. **Hanson** added that there is a high amount of uncertainty about that data point and its unknown if there were any deleterious effects on the environment. However, had that SSB level persisted, there may have been detectable effects. Also during that time, there were a number of fisheries that were overfished and in need of rebuilding which could have been buffered by a lower number of top predators at that time. Now that might not be the case. **Jones**, with that in mind, stated that we need to take all of those into consideration as we continue to discuss reference points moving forward in this meeting.

Schueller did pull out the data on the correlation of the indices to each other which was asked for earlier by **Hanson** (Figure 2). This figure affirms that the indices of abundance have a large influence on the estimate of stock size in the assessment (below). The Pearson product moment correlation coefficient value was 0.81 between the gill net and population fecundity estimates (a surrogate for SSB and the metric of SSB used in the stock assessment).

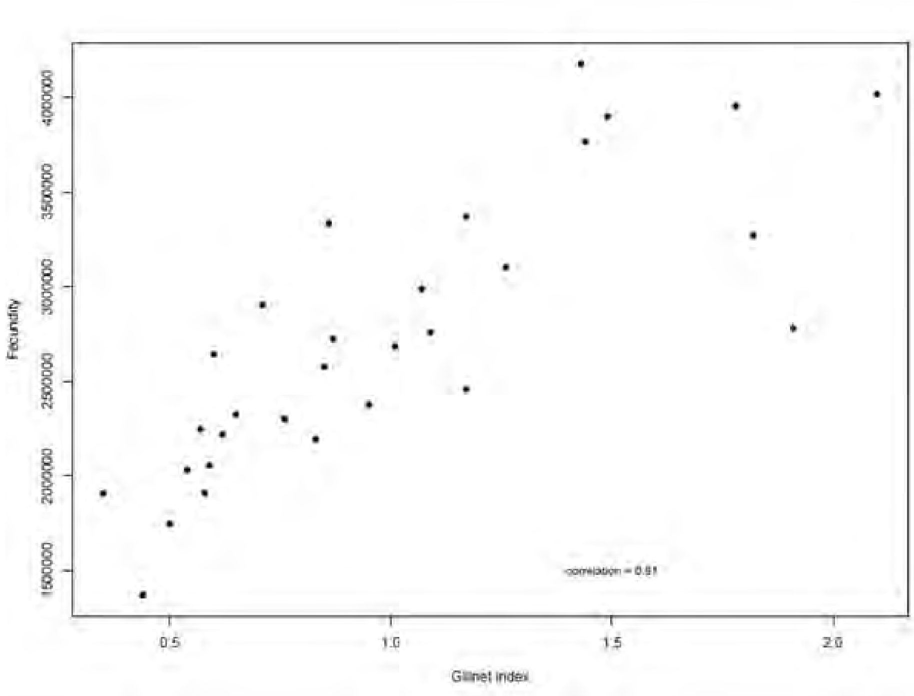


Figure 2. Plot of the gillnet index versus fecundity for 1988-2017 with the correlation = 0.81 (provided by Schueller from the SEDAR63 results).

Moving along, **Butterworth** pointed out on the SSB projection graph from the *January Paper* (Figure 2), for both historic and projected range, the future range is above the lowest point we have seen in the past. When we get into the MSE, we need to consider the consequences if one of the projections in the assessment was wrong. The SSB plot in a “worst case” scenario (Figure 2) shows what that forecast would be with ten years of continuous low recruitment (50% lower than expected each year for ten consecutive years).

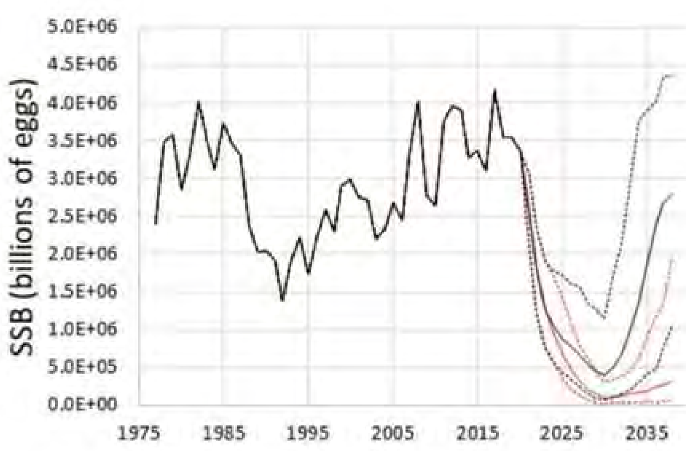


Figure 2. Historical estimates and projected 20-year median and 90%iles for a series of quantities for the **robustness test** with **10 years bad recruitment** with **Fmax=4.0**, without (red lines) and with (black lines) the management rule (Figure 5 (row three left panel) from the *January Paper*).

Figure 2 shows the results of this robustness test. The solid red line again represents if no action is taken, the population declines significantly but eventually does begin to recover at the cost of extreme reductions in harvest. The solid black line represents the response following activation of the HCR. SSB still declines but not as low and the recovery is more substantial and quicker. An open question for the group to consider is if this scenario a realistic one? It has never happened but as an extreme case, if it did, the HCR would help the population but probably not the fishery, there just would not have been enough fish to catch and they very well would be out of business. This is why it is considered an exceptional circumstance; a worst-case scenario and appropriate as a robustness test.

Butterworth noted that there are two types of robustness tests. First, what if the assessment is wrong. If we change our interpretation of the assessment, what impacts occur? Second, what if the assessment is fine but conditions in the future change from those seen historically. To be useful, robustness tests MUST be plausible. We cannot just try to break the model, it must have some basis in reality.

Hanson asked when do you turn the HCR on and how does that process work? **Butterworth** explained that the rule has a threshold, above it there is no limitation on fishery, but once you drop below a certain level, the limit is imposed based on the combination of the two indices. Looking at (Figure 3), the threshold was set at an index of 0.8.

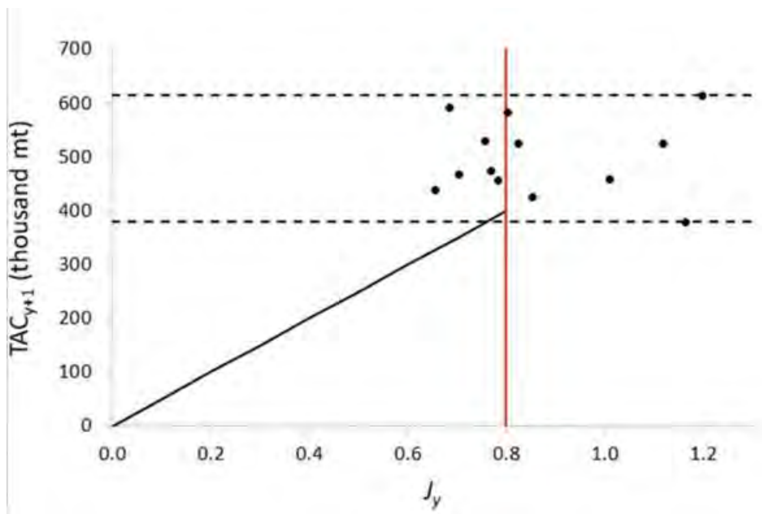


Figure 3. Illustration of the management rule for set tuning values considered in the example for which results are reported. The horizontal dash lines show the 2000-2017 minimum and maximum landing values. The historical (1999-2017) J_y vs. TAC_{y+1} are shown as black dots (Figure 2 (Top plot) from the *January Paper*).

If the combined index (J_y) ever falls below 0.8 (red line on Figure 3), the diagonal black line determines the catch limit, which depends on the index value each year. A little abundance drop below the red line results in lower TAC, a larger drop necessitates a much lower TAC. If the index continues to fall in subsequent years, the TAC will continue to drop until the index returns to 0.8 or higher. To the right of the red line, there are no limitations on the catch. The dashed horizontal lines at 380K mt and 610K mt represent the actual range of catches for the fishery over the last 13 years.

Dix wondered if the team had looked at the historical data points where the decline in SSB occurred in 1992 under a HCR. Is there any way to simulate what might have happened in and after 1992 had there been management like this? **Butterworth** returned to the future projections and explained the timing of how the HCR was activated. **Dix** understands that this is a future simulation. Is there a simulation of how the decline would or would not have happened had a HCR been available to activate historically? **Jones** explained that the rule works on future years, it cannot be retroactively be applied.

Runzel asked if this is the worst case scenario and if this has ever happened in history. **Butterworth** and **Jones** answered no, nothing has ever been as dire as that scenario but it is not unprecedented. Again, there has to be a level of plausibility in any of the scenarios they test. Admittedly, this is probably a very unrealistic series of poor recruitment however (50% annually in SSB over ten consecutive years), it is a good example of an extreme robustness test. In longer lived species, this is probably more realistic but you have many more year classes that can continue to support the SSB if you have multiple years of recruitment failure. In short lived species, that sort of scenario would be catastrophic so it is useful to include the idea even if it is taking it too far.

Runzel also asked that under the robustness test, if the recruitment scenario in Figure 2 (above) did actually play out, it would be catastrophic across the board for all species in the system, lower trophic and predators. **Butterworth** stated that we will definitely address this tomorrow. **Jones** wrapped up briefly that this work in the January paper was what drove the discussions during the RP Workshop and now we will begin to look at what has taken place since with the analytical team.

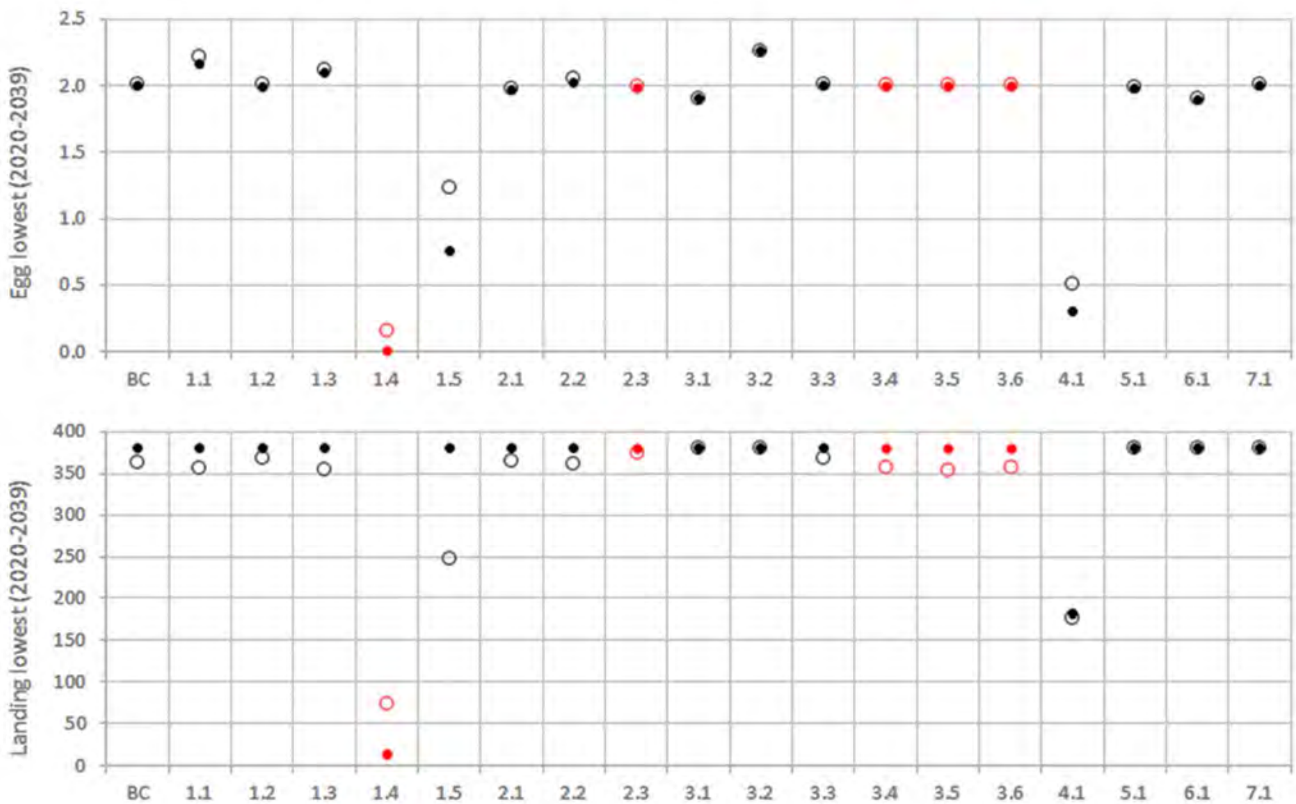
Leaf provided info on the progress made by the analysts since the RP Workshop. The team was formed of scientists that have experience in Menhaden and other fisheries and included **Drs. Genny Nesslage**, **Amy Schueller**, **David Chagaris**, and **Robert Leaf** to work with **Butterworth** and **Rademeyer** to come up with a suite of robustness scenarios They came up with the tests that were the proving ground for the HCR which we will talk about this week. The work took place through spring/summer and was wrapped up Monday and Tuesday of this week prior to this workshop. The results we will go through at this workshop are the fruit of those efforts (Table 1). **Hanson** asked if there was documentation of this work. **Leaf** explained that yes, the suite of robustness tests are included in the July paper that was sent out by **VanderKooy** ahead of this workshop. **Hanson** asked what the meeting was about that happened the past two days. **Leaf** explained that the majority of work in the July paper was finished prior to this week but these last few days were to 'fine tune' the work for preparation of this group presentation – to simplify the information so it was consolidated and more clear. **Nesslage** was on the webinar earlier this week but **Chagaris** was not available. **Jones** indicated that the table is the list of all the things that were considered by the scientists. Were the other reference points and considerations that were generated in the RP Workshop included in those discussions? Were there some things 'on the table' that are no longer following the analyst discussion? **Leaf**: yes there were a number that proved to be redundant and were consolidated or eliminated. Measures other than the indices were not really looked (such as SSB) simply because there were time constraints as far as turning this around quickly – working with only six months to complete what they did. There is no practical way to build brand new stock assessments to generate the alternatives that had been discussed but were already well correlated with the actual benchmark assessment. **Butterworth** pointed out that these exercises are very complicated and very time consuming. He reiterated that simple rules often do just as well as more complicated ones and the more complex rules often require annual assessment with multiple runs to generate the variation, so it is often just not practical.

Table 1. List of the robustness tests used in MP testing. Note that "No refitting" means that the test involves changes in the future only. Type A OMs are considered to reflect alternative plausible realities to the Base Case OM, while the plausibility of Type B OMs is low at best, but these OMs have been included more with a view to check how far the MPs considered can be "pushed" before they provide inadequate performance (Table 1 from the *July paper_fin*; *Appendix A*).

	Base Case	Robustness	No refitting	Type
1. Alternative choices for *M*				
1.1		M'(a)=1.2		A
1.2		M'(a)=M(a)*exp(-0.1(a-2))		A
1.3	Lorenzen mortality vector	M(4+)=1.67		A
1.4		M increases linearly by 40% over next 20 years	x	B
1.5		M increases linearly by 20% over next 20 years	x	A
2. Alternative catch selectivity function				
2.1	S(3) = S(4+) = 0.87	S(3) = S(4+) = 1.0		A
2.2		S(3) = S(4+) = 0.74		A
2.3	S(1) in future as estimated in past	S(1) in future, double that estimated in the past	x	B
3. Indices				
3.1	Linear relationship to abundance: I = q*B	sqrt relationship to abundance I =q*sqrt(B)		A
3.2	Weighting: 4:1 gillnet to seine	Weighting: 1:1 gillnet to seine		A
3.3		Observation error = 0.2	x	A
3.4	Observation error = 0.11	Observation error = 0.3	x	B
3.5		Observation error = 0.5	x	B
3.6	Flat 2+ gillnet selectivity in the future	Increasing 2+ sel. slope over the next 20 years (to 0.4 age 4 in 20yrs)	x	B
4. Period of future poor recruitment				
4.1	Future rec. drawn at random from past values	Five (2020-2024) years of bad recruitments (50%)	x	A
5. Alternative stock-recruitment function				
5.1	Hockey-stick, hinge-point=1.8 billion eggs	Hockey-stick, hinge-point=2.2 billion eggs	x	A
6. Under-reporting of future catches (which is not noticed)				
6.1	Future catches=TAC	Future catches = 1.1TAC (presence of these IUU catches is not realised)	x	A
7. Maximal possible fishing mortality				
7.1	Fmax for projections = 1.05*Fmax historical	Fmax for projections = 1.20*Fmax historical	x	A

Hanson indicated that we talked at length about the uncertainty in those indices and the reliability of the fishery independent data since it is not a survey targeted at Menhaden. **Leaf** noted that whatever control parameter we agreed upon needed to be based on data we already have. It would be unrealistic to develop a new survey that would be used for this, financially and practically. **Hanson** pointed out that there are other ways to ground-truth, perhaps other collaborating evidence in smaller, or shorter duration surveys. Is there other data out there that could be used to corroborate? **Schueller** stated that if there is something out there that we are not aware of, and that the assessment team would want to know that. Over the evolution of this assessment, they have tried to leave no stone unturned regarding fishery independent data sources. The indices here are the best available science at this moment in time. The assessment is predicated on that.

Butterworth summarized the Base Case (BC) Table 1 which explains all the potential metrics that were explored and discussed by the analysts for the MSE. The BC situation we will present tomorrow demonstrates the current status of the population and how it may track moving into the future (20 years). **Butterworth** pointed to the right hand column which shows the options such that Type A were more plausible simulations and Type B alternatives were less plausible but intended to see if, under these unlikely conditions, the conclusions about which HCRs are acceptable would change. **Jones** reminded that one of the issues we discussed at the RP Workshop is in 2.3 in Table 1. The discussion was around the selectivity in the catch changing over time. If there was a shift to younger fish by the fishery in the future, what would be the result? That question was explored in the robustness test 2.3. This is an example of how the analytical team kicked the model around to be sure that the model is resilient to a number of uncertainties about the fishery and population. This concluded our agenda for today; the remaining time will be used to allow the analysts to fine tune their presentations for tomorrow's meeting. **Butterworth** asked the group to do some homework and look at Figure 4 (Figure 3 in the *July paper_fin; Appendix A*) for tomorrow. This figure displays key results of the robustness tests that were listed in the table above. Looking at the BC on the left and going across, there were very few differences from the Base Case. The three robustness tests that are different are the ones we



will focus in on tomorrow.

Figure 4. Median lowest egg production and landing values over the 2020-2039 projection period for each of the Base Case and Robustness test OMs without (full circles) and with (open circles) the Baseline MP. 'Type B' OMs are shown in red (Figure 3 from the *July paper_fin; Appendix A*).

Donaldson had some closing remarks for the group and thanked everyone for attending this important workshop. The Commission has a long history of working with the industry and stakeholders in the Gulf. We provide a forum for a wide variety of issues including Menhaden and this workshop is a good example of how we facilitate moving towards solutions. Thanks again.

The Wednesday workshop adjourned at 4:12 p.m.

Thursday, July 18 (8:00 am – 5:00 pm)
Jones began with a review from the previous day. He recapped the series of changes that were made to the base model, as discussed yesterday. Today we work through the results of the various robustness tests.

Butterworth brought attention to the *July paper_fin; Appendix A* (*Extensions of the Application of MSE to Gulf Menhaden*) which **VanderKooy** had circulated before the meeting. The third bullet point on page 2 of this paper explains how the operating model handles the situation where an unrealistic harvest might result from an improbably high F, first by lowering selectivity on age-2s, and then putting a cap on F. **Hanson** asked about the amount of harvest that could not be maintained. **Butterworth** is referring to the fishery maintaining the historic levels between 400K and 600K mt. **Dix** asked about the uncertainty when ageing and raised concern over the quality of the age data. **Schueller** mentioned that they had done a series of internal checks to ensure consistency of age estimates and the assessment and ageing group assumed that it is consistent. They went back and had Ethel reread several decades worth of scales to be sure the results were consistent, and they were. **Jones** asked **Schueller** about the ability to assign Menhaden ages relative to other species. **Schueller** assured the group that the ageing of Atlantic Menhaden is very reliable compared to all species they work on therefore there is a high level of confidence in the process. Per **Jones** ageing is a minor issue in regard to assessment of Menhaden.

Butterworth also noted the final bullet on Page 2 which explains that we assume a lognormal random observation error for the abundance index. We also assume autocorrelation for the seine index: this implies that if the abundance index error is positive in one year, then is it likely to be positive again in the next year. We did not assume autocorrelation in the gill net index, because the data did not show evidence of a positive correlation.

Jones wanted to check in with the industry to determine if in fact the fishery could actually shift the selectivity that much if age-2s were fewer. **Swindell** indicated that there is no way to effectively fish for fish an age group. They generally target larger fish and depending on where you are in the range of the species, you will get one or the other and they are generally not mixed that you could actually 'target' younger fish. Age-1s are generally dominant in the eastern Gulf. **Butterworth** is not sure that selectivity is understood – it is not targeting. It is relative to the numbers of age groups that are there and availability – selectivity indicates how many of that group CAN be caught by the gear, not how the industry picks fish to target. **Landry** stated that in terms of targeting, the spotters are talented and can usually determine the age class from air. Age-2s and mature age-1s look similar therefore it would be impossible to target one over the other. **Butterworth's** take is that if it is assumed that the industry can be more efficient in how they harvest, this provides an additional level of conservatism, in that they MIGHT be able to hit the resource harder than they actually do.

Butterworth revisited Table 1 (above) from the July paper and looked at each of the robustness alternatives. The first three alternatives under 1 (Natural Mortality) generally address changes to the assessment model assumptions. Alternatives 1.4 and 1.5 reflect ecosystem-related robustness tests, based on what **Chagaris** provided. These are ones where, in the future, natural mortality increases, possibly due to increased predators (although other mechanisms such as environmental change are also possible). In 2 (Alternative Catch Selectivity), 2.1 and 2.2 deal with adjustments to the assumed selectivity for older fish. 2.3 explores a future where selectivity on age-1s doubles relative to estimated levels in the past. **Hanson** asked, alternative 1.4 is tagged with a Type B which means it is unrealistic but is it? **Butterworth** – could this happen in a few years on occasion? Yes. However, 1.4 is an increase in EVERY year over the 20 and so it is less likely to be continuous. It was the opinion of the modeling group that this scenario was quite unlikely.

Hanson asked about observation error and what that actually is. **Butterworth** referred to Figure 5 and looked at the panels in the top rows. This is the result of the assessment showing how the model fits the observed seine and gill net index. When we talk about typical error being about 10%, we look at the standard deviation of the residuals between observed and predicted (modeled) index values. **Hanson**: "I understand observation error but is there also uncertainty around the indices where the point was generated to begin with?" **Butterworth** explained that we are comparing the model estimate to the index itself. There is error in the sampling. There is error in the recruitment estimate because it only is a point in time which may or may not adequately represent the true recruitment. So what we are doing is taking ALL that error and combining them as we project into the future and that gives the wider range of potential realities in the projections.

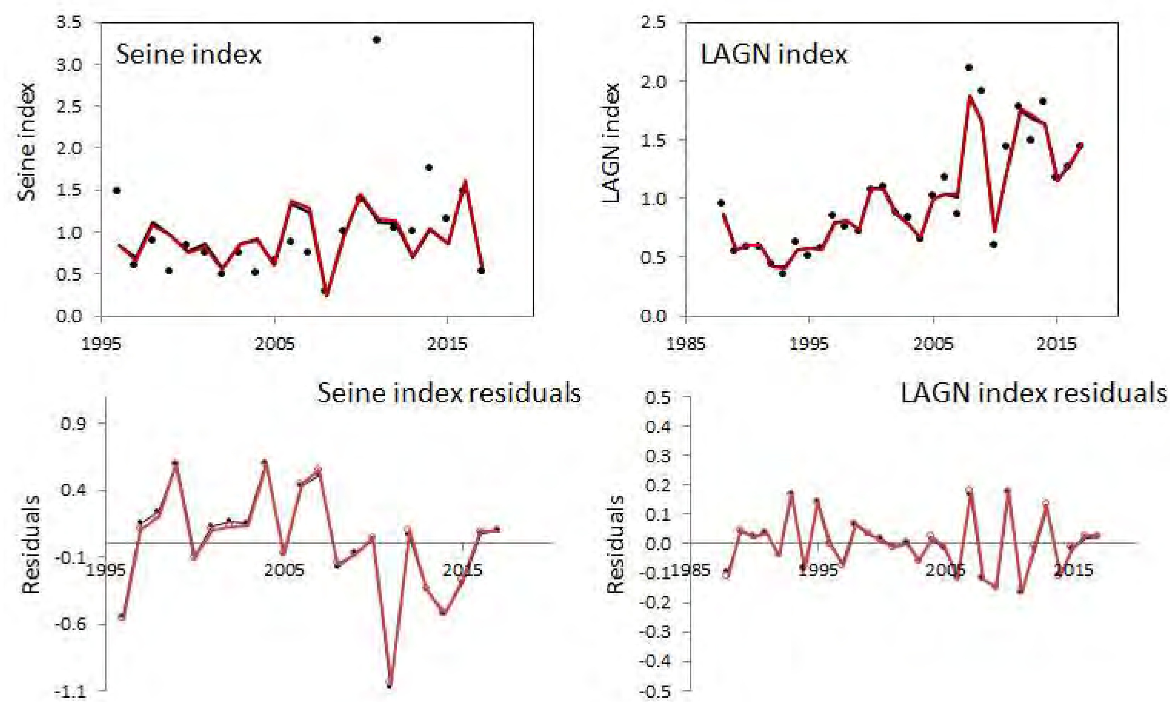


Figure 5. Assessment results for the Base Case (black lines) and Robustness test 1.1 (M = 1.2) OMs (Figure D1a in the *July paper_fin; Appendix A*).

Butterworth returned to the summary of robustness scenarios, Table 1. We spent some time discussing, yesterday, the Robustness test 4 (Period of Future Poor Recruitment). **Butterworth** brought up Figure 6 and discussed historic recruitments with the hockey-stick fit to give more conservative stock-recruit relationship which addresses Alternative 5 in Table 1. There is no indication of trend in the recruitment data. Extrapolation to lower SSB levels is the issue we need to address when recruitment varies so widely. At some point it has to drop off however so without many SSB observations below 2.0 billions of eggs, we provide a straight line that is a pessimistic assumption you can make if there is a serious reduction in SSB. The 2.2 billion break-point is even more conservative (recruitment starts to fall off at even higher SSB levels).

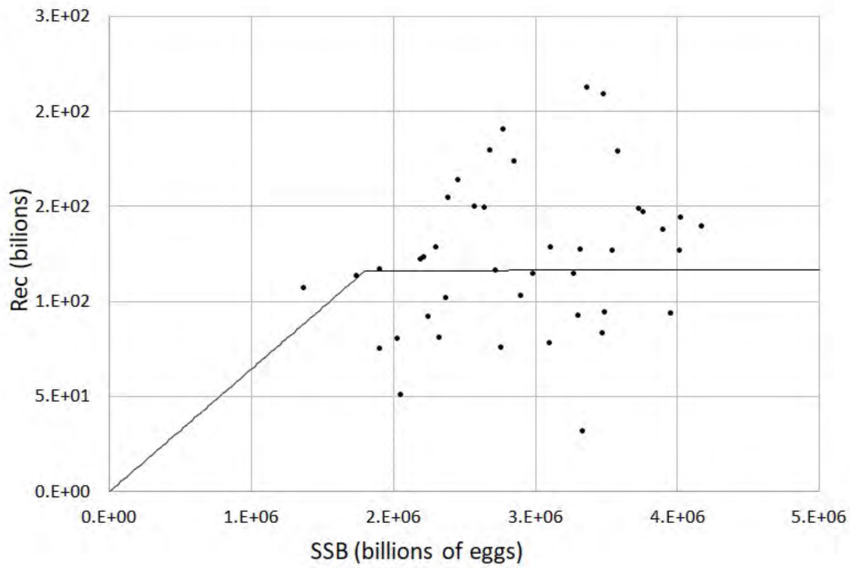


Figure 6. Hockey-stick stock recruitment curve for Gulf Menhaden which is used to compute projected recruitment. The data points are those estimated in the BAM Base Model (Figure 1 in the *July paper_fin; Appendix A*).

Butterworth returned to Table 1 to look at Alternative 6 regarding under-reporting of catches should a restriction be put in place. At times, there could be a temptation or pressure in a valuable or limited fishery to under-report in that situation so the test provides some security in the event it would become an issue. **Hanson** asked if the analyst team discussed what would happen if they overestimated what they caught. **Butterworth** stated that this could be tested but from a conservation view, these would not affect the outcome. No one on the analyst team even raised it but if we test it, we already have a good idea of what the results would be. **Leaf** stated that all robustness scenarios are examining uncertainty, focusing on worst-case/pessimistic scenarios. The robustness tests do not reflect what the modelers think is likely. They are intentionally pessimistic and assume that the resource is NOT as resilient as we think it is.

Finally, **Butterworth** reviewed the inclusion of Alternative 7 in the robustness tests (Table 1) which spreads the selectivity as discussed earlier.

Next, we need to look at the relative importance of each of these scenarios. Yesterday's homework was for the participants to look at Figure 4 (above) which shows the results of each robustness test as well as with and without the HCR in place. Only, three of the tests fall out as having a large effect relative to the first column, which is the Base Case: Robustness tests 1.4, 1.5 (increases in natural mortality), and 4.1 (five years of poor recruitment). From the resource point of view, we do not want to see egg production to go down too far and those three scenarios had substantial declines looking out 20 years from now. From the industry point of view, it important to have at least some consistency over time even if it requires a slight restriction of harvest. It is better to act sooner with a slight decline in harvest than wait too long and risk the ability to stay in business.

Runzel noted that 1.4 and 1.5 were increases in natural mortality due to increased predation but have the implications of natural mortality increases due to natural disasters and sea surface temperatures been considered in the robustness tests. **Butterworth** stated that this might be something worth looking at, not so much climate change which is more a trend and is arguably represented by 1.4 and 1.5, but maybe as a punctuated event (stochastic event) such as the oil spill or something similar. **Butterworth** agreed that some tests of this nature should be done and perhaps we could look at this following this workshop. **Runzel** also asked how that sea surface temperature change could affect the HCR and the rebuilding of the stock into the future. **Butterworth** – this is another good one that should be added to our work in the next round to essentially see if the carrying capacity of the resource goes down. These are ecosystem effects in the same way as natural mortality, which was the obvious one to us on the team. Does a slow change provide enough signal for the HCR to respond?

Hanson pointed out in the M alternatives 1.4 and 1.5, a 40% increase in predation may actually be a plausible scenario despite it being listed as a Type B in the table by the analysts when you add in those other compounding effects. Where is the boundary cutoff for M? **Butterworth** will address this in the results which are still coming. We are trying to mimic a number of potential nasty events but we are not necessarily trying to break the model. We need to balance out the bad scenarios with the reasonable responses based upon what HAS happened in history. **Jones** stated that one important layer that is not captured might be some of those shock events.

Schueller mentioned that in the relatively recent past, we actually have had two shock events, Katrina and BP oil spill and when those events occurred, it limited fishing so it reduced effort. These were both considered in the assessment and were talked about in the report. These are in the scope of what was considered as future environmental conditions already. **Hanson** thinks that hypoxia and natural mortality, sea level rise, red tide, should be part of the encapsulation of what factors are potentially affecting recruitment and need to be included if they are not already. Figure 5 (above) is treating all those alternatives individually and not in combination. Is there a possibility to combine them? **Butterworth** stated that **Hanson** is correct, these are all single perturbations and perhaps combined effects could be considered if we go in to another phase of analysis, but these that are not making an impact individually and would not likely make a great impact in combination so we need to use common sense with selecting combinations. **Hanson** suggested that a combination of more correlated effects could be explored, not all of them. **Kuttel** reminded the group that we need remember that the Base Case is the most plausible scenario. Some big events have already occurred and are implicit in the data which this model is based. **Adriance** also argued to **Schueller's** point that hypoxia and natural events are woven into the results already and that hypoxia is less likely to affect surface dwelling pelagic fish which generally move. **Cufone** stated that NOAA predicts these events will get more intense and frequent in the future. **Jones** wrote some ideas for additional analyses on the flipchart based on the comments so far but moving forward in the results we may better define some of these. **Hanson** also pointed out the water diversion projects that are happening and as these continue or become more frequent due to climate and weather pattern changes, this should not be ignored and should be considered. **Cufone** agreed and noted that there are lots of examples of cascading affects in other fisheries from perturbations to the environment. The models are only as good as the inputs going in it. **Adriance** stated that the modeling **Hanson** was mentioning is taking place with Coastal

Protection and Restoration Authority (CPRA) and the LDWF. **Cufone** thinks inclusion of the CPRA in these discussions would be valuable. **Jones** summarized that we need to look at what aspects of possible environmental change might affect recruitment or survival of Menhaden and whether there are things that the robustness tests we are considering would not necessarily capture.

Break

Butterworth continued to explain making projections (see Appendix A in the *July paper_fin*) which. Appendix A of the *July paper_fin* explains the steps in the process, which was very similar to the process in the January paper. Appendix B of the *July paper_fin* presents the rule used to set the TAC. There are actually three inputs. Once the simulated index falls below the J_y threshold, the control rule, comes into play. To apply the HCR, you must have all the fishery-independent data available to derive the J_y threshold. The p input is the number of previous years you are averaging together. The more years you include, the less noise but with more years you might average out a more recent signal. The catch (TAC) needs to be determined no more than six months after data are finalized so the process MUST be timely for it to work properly. Fortunately, the data are available prior to the start of the next season but the state agencies must get the data finalized as soon as possible. So in summary, in this formula there are only three values to select; threshold value (which has an element of risk), value of gamma (the risk prone or risk averse), and the value of p. **Leaf** interjected that “risk prone” means risk to the stock or resource and not to the fishery and that the MSE evaluates both risks simultaneously. So the HCR (Figure 3 above) provides the threshold index of $J_y = 0.8$; higher index values require no action, lower values result in a reduction in allowable catch. **Butterworth** noted that the rule is applied annually, but that there are other variants that can be added. **Hanson** thinks of it more of an accountability measure on a year by year basis so that if the index goes up next year, there is no TAC requirement, so it can switch on and switch off annually.

Butterworth and the analysts had found a glitch in Figure 3 and revised it in a new set of plots which **Rademeyer** had sent overnight. **VanderKooy** distributed the new material electronically to the group. Butterworth used Figure 7 until everyone had received the additional files. The graphs show forecasted landings and SSB or egg production over the next 20 years, and include the assessment estimates from 1977 to present. The spaghetti represent different possible future trajectories, given uncertainty, for p=2 and J=0.9. For this scenario in the Base Case, the rule came into play only 10% of the time. Under robustness tests 1.5 and 4.1 (increased mortality and poor recruitment respectively), the rule does come into play frequently.

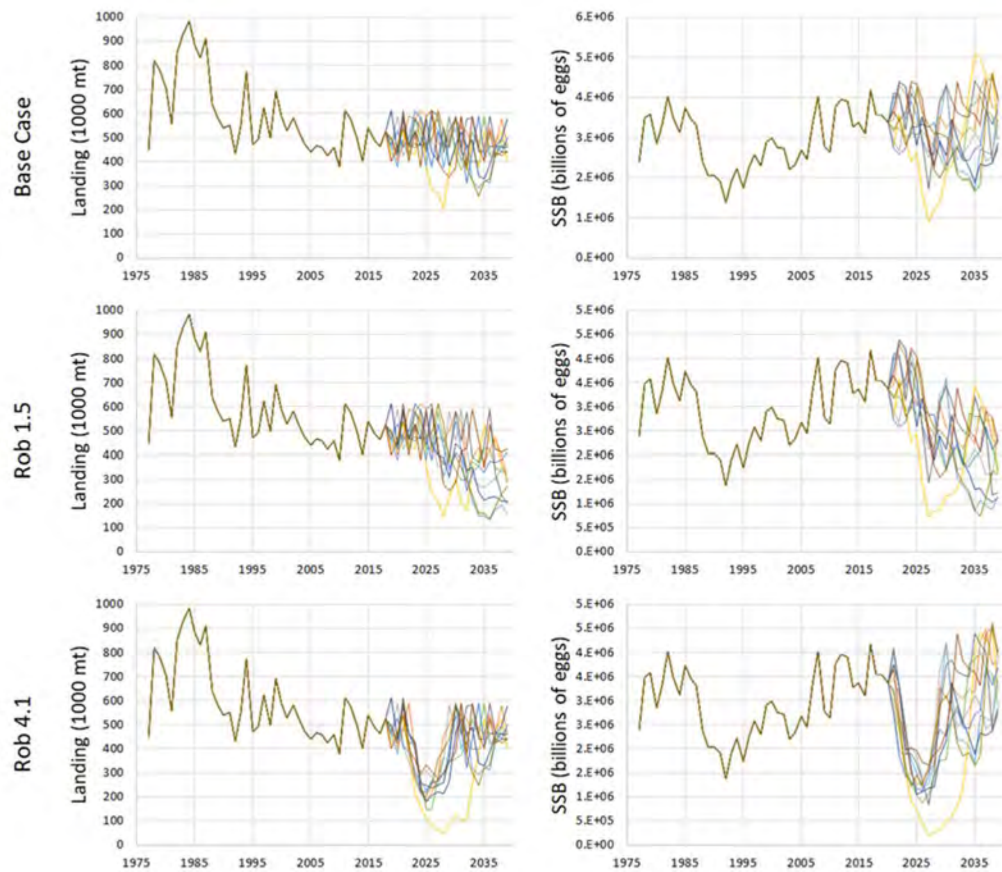


Figure 7. Worm plots for projected landings and SSB for the advocated MP variant with p = 2 for $J_{threshold} = 0.9$, for the Base Case and Robustness tests 1.5 and 4.1 (Figure 7 from the *July paper_fin; Appendix A*).

Butterworth also referred to Figure 8 and pointed to the left panels in the first and second rows, which are the same as above for the Base Case, but with the median and upper and lower 10% percentiles rather than a sample of individual simulations. The red line is without the HCR; the black is with it in place. **Butterworth** noted that the lower 10th percentile of the projections for SSB, even without the HCR in place, does not fall below the lowest level in 1992. The pie chart in Figure 8 indicates how many times the model worked correctly and incorrectly giving either false positives and negatives (red and yellow). The green and blue wedges are where the HCR worked properly. You want it to work properly and in our example, it is correct 84% of the time, the false negatives are when it should have and false positives applied the rule when it was not necessary

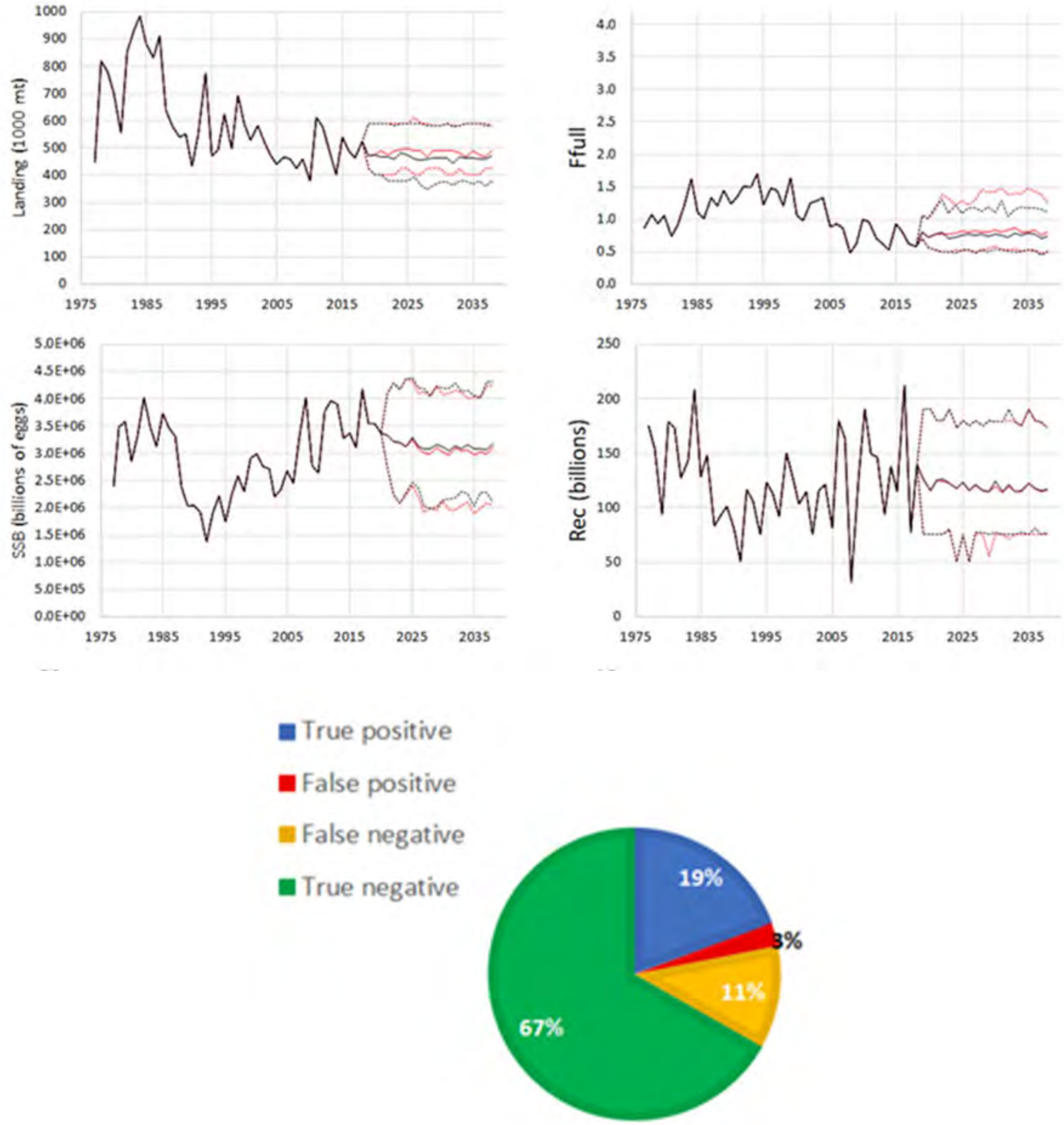


Figure 8. Historical estimates and projected 20-year median with 10%- and 90%-iles for a series of quantities for the Base Case OM, without (red lines) and with (black lines) the advocated management rule (Figure 8a from the *July paper_fin; Appendix A*).

Finally, referring to Table 2, **Butterworth** showed about 20 different indicator values for each of the most relevant robustness tests (Base Case, higher mortality, and low recruitment). We will be focusing in on five of the rows; 1) average landings over the period, 2) the lowest landings over the period, 3) lowest egg value, 4) catch variance as AAV (average annual variability) over the period, and 5) the fraction of the years the rule was applied.

	Base Case									Robustness test 1.5									Robustness test 4.1									
	No rule			MP_2_0.9			MP_2_0.9_quad			No rule			MP_2_0.9			MP_2_0.9_quad			No rule			MP_2_0.9			MP_2
	Median	10	90	Median	10	90	Median	10	90	Median	10	90	Median	10	90	Median	10	90	Median	10	90	Median	10	90	Median
Related to catch																									
Average landing 2020-2039	494.7	479.2	517.5	472.3	440.8	502.4	479.8	445.4	509.6	363.3	288.9	432.8	414.0	364.5	455.0	417.4	356.3	452.6	371.9	169.9	500.7	395.9	323.8	439.5	393.9
Av landing no rule	494.7	479.2	517.5	496.3	476.4	517.4	495.0	476.4	519.4	363.3	288.9	432.8	501.1	465.7	530.0	500.5	462.5	527.8	371.9	169.9	500.7	495.4	470.3	526.7	494.4
Av landing with rule	-			390.0	330.1	426.7	415.6	325.4	496.6	-			347.0	271.5	388.5	344.8	255.1	406.6	-			290.9	221.3	352.8	277.2
Lowest landing (2020-2039)	379.9	379.9	425.6	347.5	261.6	397.2	322.2	190.0	400.7	14.6	2.8	69.3	236.1	140.2	318.3	167.9	73.2	265.3	181.9	41.0	379.9	174.7	84.0	254.2	97.3
2020 landing	473.7	400.7	590.8	473.7	400.7	590.8	473.7	400.7	590.8	473.7	400.7	590.8	473.7	400.7	590.8	473.7	400.7	590.8	473.7	400.7	590.8	473.7	400.7	590.8	473.7
Related to abundance																									
Egg(2020)	3.32	2.73	4.10	3.32	2.73	4.10	3.32	2.73	4.10	3.32	2.73	4.10	3.32	2.73	4.10	3.32	2.73	4.10	3.32	2.73	4.10	3.32	2.73	4.10	3.32
Egg(2040)	3.17	2.21	4.04	3.20	2.37	4.07	3.20	2.37	4.05	0.01	0.00	0.07	1.63	0.89	2.39	1.76	1.01	2.40	0.59	0.10	3.80	3.19	2.37	4.07	3.20
Egg lowest (2020-2039)	2.00	1.44	2.38	2.02	1.60	2.40	2.01	1.52	2.39	0.01	0.00	0.07	1.32	0.77	1.71	1.31	0.88	1.67	0.30	0.05	1.06	1.00	0.48	1.50	1.00
Prob Egg(2040) lowest	6			5			5			100	100	100	24			19			9	9	9	0			0
Related to catch variability																									
AAV 2020-2039	0.15	0.12	0.19	0.17	0.12	0.21	0.18	0.12	0.25	0.22	0.17	0.28	0.17	0.12	0.25	0.25	0.17	0.35	0.20	0.14	0.27	0.21	0.16	0.27	0.30
AAV with rule	-			0.16	0.05	0.24	0.20	0.05	0.43	-			0.17	0.12	0.24	0.29	0.19	0.46	-			0.23	0.15	0.35	0.41
Other																			0						
Fraction years rule applied	0.00	0.00	0.00	0.20	0.05	0.40	0.20	0.05	0.40	0.00	0.00	0.00	0.60	0.35	0.70	0.60	0.35	0.70	0.00	0.00	0.00	0.50	0.35	0.65	0.50
True negative	0.0			19.4			20.8			0.0			58.0			58.4			0.0			48.7			47.1
Fals negative	0.0			2.6			2.3			0.0			1.5			1.4			0.0			1.9			2.1
False positive	0.0			11.4			11.6			0.0			11.6			11.9			0.0			10.0			10.3
True positive	100.0			66.8			65.4			100.0			29.0			28.4			100.0			39.5			40.6
Prob rule in 2020	0			0			0			0			0	0	0	0	0	0	0			0	0	0	0
Fraction years Hit Fmax	0	0	0.15	0	0	0.05	0	0	0.05	0.6	0.4	0.75	0.05	0	0.1	0.05	0	0.15	0.85	0.25	0.9	0.05	0	0.15	0.05
Hit Fmax, landings not taken	0	0	0	0	0	0	0	0	0	0.4	0.25	0.55	0	0	0	0	0	0	0.65	0	0.85	0	0	0	0

Table 2. Performance statistics for the Base Case OM and Robustness tests 1.5 and 4.1 without the management rule ("No rule"), with the advocated MP ("MP_2_0.9") and with the quadratic MP ("MP_2_0.9_quad") (Table 3 from the *July paper_fin*).

Break

Butterworth asked everyone to look at the file titled *Further resultsA_18July*. Figure 9 (below) shows the difference in tuning using the *p* values of 2 or 3 (the number of years over which you average the index, *J*). The curve for *p*=3 is smoother and less likely to require frequent turning off and on of the HCR but has the potential to delay response to a change in the resource. The industry would prefer *p*=3 as it provides greater consistency among years.

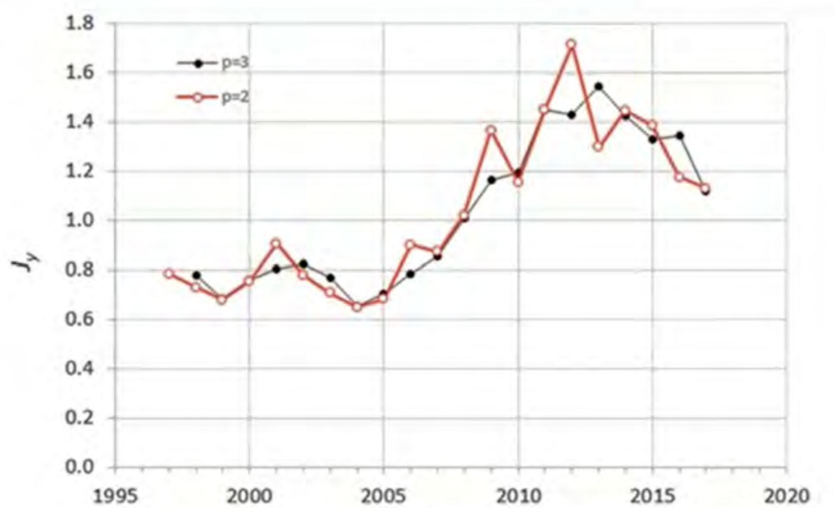


Figure 9: Historical combined index J_y values for *p*=2 and *p*=3 (Figure 2 from the *Further resultsA_18July*).

Butterworth presented the modified HCR figure with additional tuning (Figure 10). There are multiple levels of risk tuning around the two proposed index thresholds. The July paper suggested J_y=0.8 originally, but results here bracket this level, at 0.7 and 0.9. The thresholds are shown along with the associated TAC (diagonal line) that would result if the index falls below the threshold. For each threshold there are alternatives (**Butterworth** called them tunings) representing different levels of risk (to the resource) tolerance, which allow managers to ease or increase the restriction on the TAC depending on how much risk they are willing to accept. A lower tuning value (dotted lines) implies a greater risk tolerance. A higher tuning value (dashed lines) implies greater risk aversion. For the 0.7 threshold, there is no dashed line for 1.25 (most risk averse) because the rule would require the TAC to be zero if the index fell below 0.7. For index thresholds above 0.9, it does nothing for the resource but would have negative consequences for the industry. Index thresholds lower than 0.7 were judged by the modelers to be too risky for the resource. The black dots are the last ten years of landings, all of which are above the threshold and thus would not necessitate TAC reductions. The open circles were earlier (1997-2006) and would have activated the HCR depending on which threshold was selected. The results for p=3 are similar for a threshold of 0.9 but for a threshold of 0.7 even the “risk-neutral” tuning results in fishery closure if the threshold is exceeded.

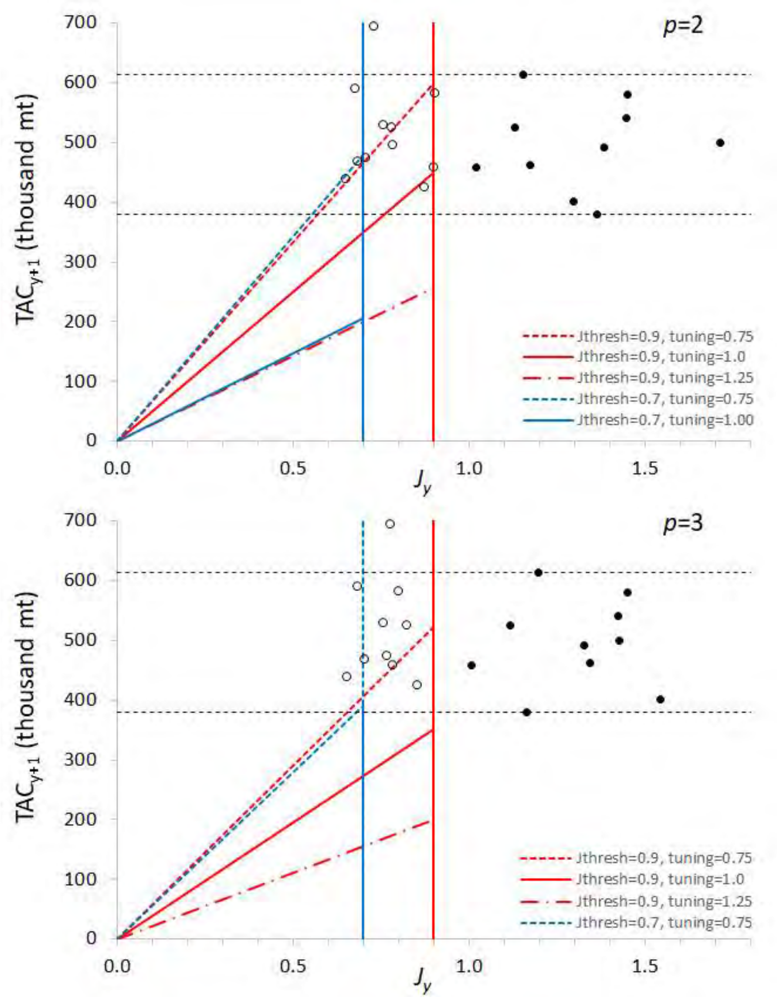


Figure 10. Illustration of the MP variants' rule for two value of the $J_{threshold}$ control parameter, with the γ control parameter value tuned so that the median lowest SSB for Robustness test 4.1 (t) is equal to 0.75, 1.0 and 1.25. In some cases, the target could not be reached ($J_{threshold}$=0.7, t=1.25 for p=2 and $J_{threshold}$=0.7, t=1.00 and 1.25 for p=3). The horizontal dash lines show the 2009-2018 minimum and maximum landing values. The last 10 years' historical (2008-2017) J_y vs TAC_{y+1} are shown as black dots (Figure 1 from the *Further resultsA_18July*).

Looking at this scenario, **Butterworth** explained how often you would apply the HCR using the three different tuning conditions (0.75, 1.0, and 1.25) for the slope on the TAC in Figure 10. In Table 3, you can see the fraction of years (or change to a percentage) of times in the 20 year forecast projections for each value. In the first column and row (the lowest risk factor at the lowest index trigger point) you would expect to enact the HCR one time in 20 years or 5% of the time. With the threshold at 0.9, it would be applied four times in 20 or 25% of the time. Once we have a problem, the rule would come into play more often as well until the resource returned to a normal level.

Table 3. Median fraction of years the rule is applied for the Base Case and Robustness tests 1.5 and 4.1 for a series of MP variants (Table 2 from *Further resultsA_18July*).

Base Case		t=0.75	t=1.0	t=1.25
p=2	Jthresh=0.7	0.05	0.05	x
	Jthresh=0.9	0.25	0.20	0.15
p=3	Jthresh=0.7	0.00	x	x
	Jthresh=0.9	0.20	0.15	0.15

Rob 1.5		t=0.75	t=1.0	t=1.25
p=2	Jthresh=0.7	0.35	0.25	x
	Jthresh=0.9	0.63	0.60	0.45
p=3	Jthresh=0.7	0.30	x	x
	Jthresh=0.9	0.60	0.50	0.43

Rob 4.1		t=0.75	t=1.0	t=1.25
p=2	Jthresh=0.7	0.40	0.30	x
	Jthresh=0.9	0.60	0.50	0.40
p=3	Jthresh=0.7	0.40	x	x
	Jthresh=0.9	0.55	0.45	0.40

Butterworth further explained Figure 11. These figures show median values of landings and SSB for various rules, no rule (red), and a case with no harvest (green). The three rows show the Base Case and two robustness cases. For the Base Case there is very little difference among HCRs. For the robustness cases, you get recovery compared to the red line (no HCR) – some are slower but all recover and are better than no rule (red).

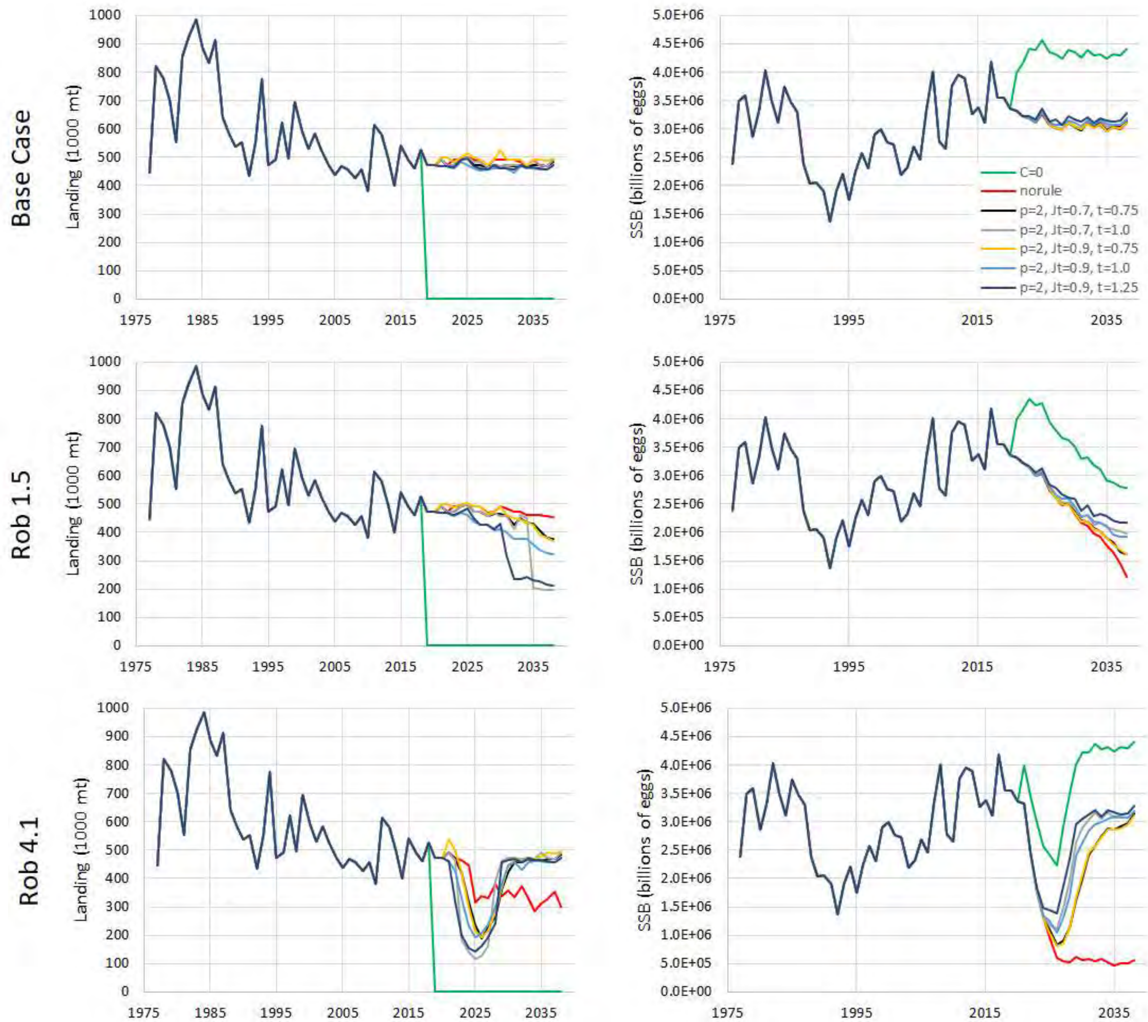


Figure 11. Historical estimates and projected 20-year for a series of quantities for the **Base Case** OM and **Robustness tests, 1.5 and 4.1**, with no future catch (green lines), without (red lines) and with a series of the MP variants with p=2 (Figure 4a from *Further resultsA_18July*).

Butterworth now reviewed Figure 12. Each dot represents the results of an HCR option (index threshold, tuning or risk, and *p* value or years averaged), including no HCR (no rule) and no fishery (C=0) for comparison. Four performance measures are shown (SSB: top row; lowest landing: 2nd row; average landing: 3rd row; catch variation: 4th row). Across the x-axis are the Base Case and the two robustness tests (1.5 and 4.). The whiskers represent the 90th percentile range of values.

The best outcomes for the resource appear to result from the 6th and 7th options (blue and dirty red). The best outcomes for the fishery industry appear to result from the 3rd and 5th options (yellow and green). This is how we need to review the options and try to achieve a balance between the resource and the industry.

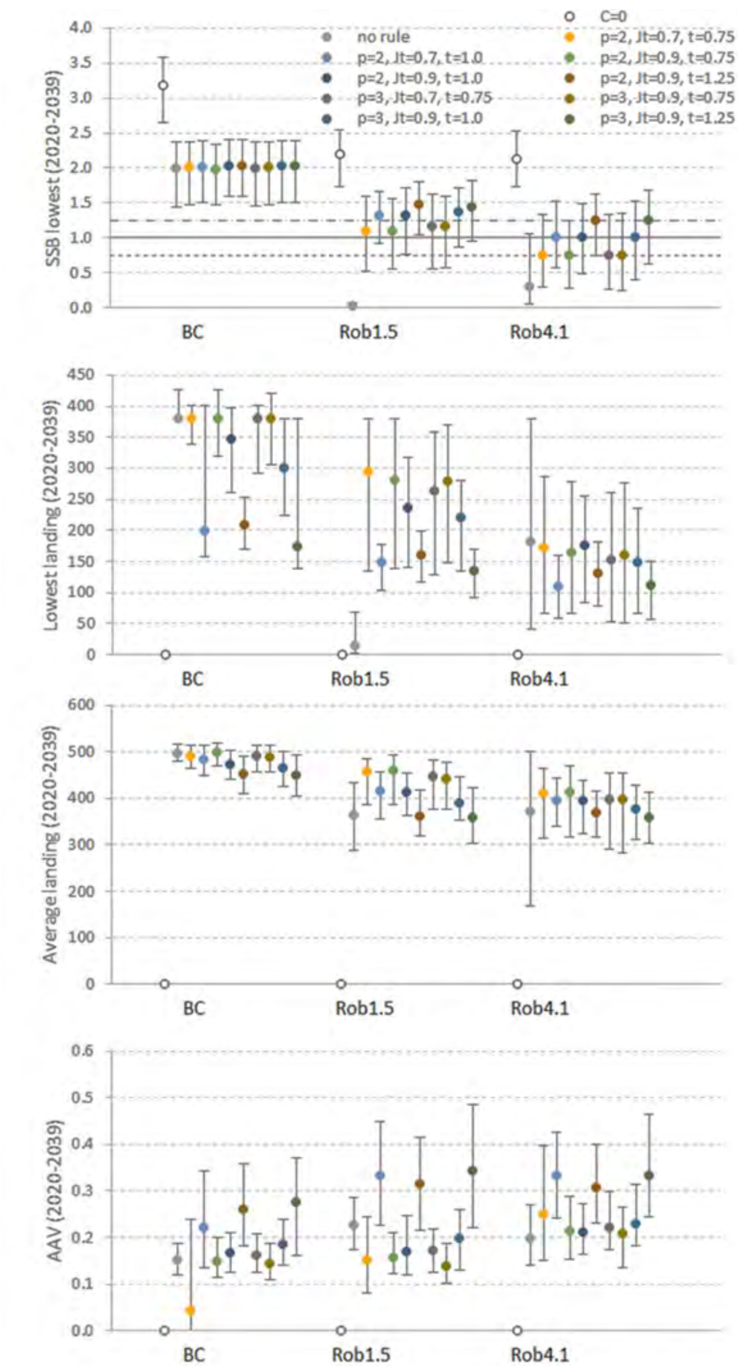


Figure 12. Performance statistics for no landings, no harvest control rule, and MP variants (Figure 3 in the *Further resultsA_18July*).

From a resource risk perspective, we could take green off the table because the blue option is just as good for the resource as is green. Now as we look at the tuning levels so that the slopes of the line respond in the Robustness test 4.1 (low recruitment), it is defined by having the lowest level that the SSB for most of the HCR options. Each tuning level moves the median lower or higher depending on the resource risk. **Butterworth** asked everyone to look over these options in Figure 12 and look at where each person's comfort level is with the various HCRs. **Hanson** asked to have an explanation about the 1.0 line on the SSB plot. **Butterworth** showed the top right panel in Figure 11 (above) the Base Case includes the SSB results down to 1.5 historically. In the bottom right panel, the low recruitment test has the projection drop almost to 1.0. Tuning drives the median of the drop higher or lower compared to the red line (no HCR) vs green line (no fishing – total closure).

Jones stated that after lunch the group would discuss the range of management procedures which have been presented here and evolve into discussion about the results that we have versus what more we need to do, to get where we can reach some consensus. These latest results were a lot to take in, but now we will move into what it all means.

Lunch Break

Jones reported that another document was sent by **VanderKooy** entitled *Trade_off_plots_v2; Appendix C* and that we would be looking at these graphs next. **Butterworth** explained that these graphs were based on the previous Figure 2 (above) but plotted on the same scale as Figure 13 in order to more easily compare performance among options, and identify tradeoffs between landings metrics on the y-axis and resource (SSB) metrics on the x-axis.

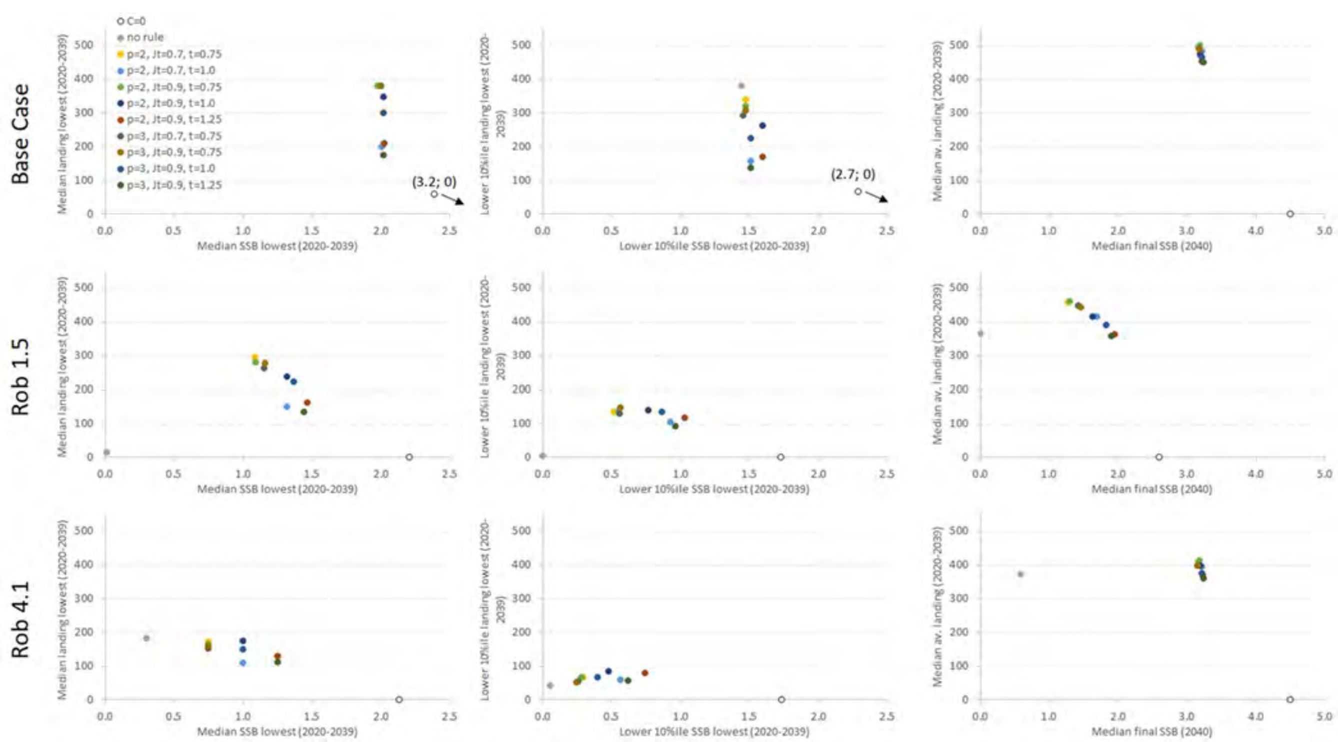


Figure 13. Trade-off plots for the Base Case, and Robustness tests 1.5 and 4.1 under no catch, no rule and a series of MP variants (Figure 1 from the *Trade_off_plots_v2; Appendix C*).

Better outcomes for the resource are to the right along the x-axis and better outcomes for the fishery are further up on the y-axis. When the options stack directly over each other as for the Base Case (especially the upper left), there is no difference in outcome among options for the resource. The two robustness tests show more of a tradeoff. In the bottom right, when you have decreased recruitment, the only option that really hurts the resource is NOT having a rule. The rest are about the same. The reason the options tend to cluster is because that is exactly how we assigned the risk or tuning (0.75, 1.00, and 1.25) as seen in the median SSB plots.

Runzel would still like to see simulations that include shock events and sea temp reflected in these charts. **Butterworth** will be sure that it would be incorporated into the future simulations. **Jones** suggested everyone reserve time to address this specifically and make sure any additional simulations can be agreed upon for the next round of results. **Butterworth** stated that a shock event is like a period of poor recruitment. It reacts the same. The impact is what you see on the lowest row of panels. Also declining habitat or environmental conditions would most likely result in lower survival or increased natural mortality so this scenario is also sort of included here already – at least the effect is. At the end of the day, these environmental change scenarios will either kill off what you have already (mortality) or have an impact on what is coming in (recruitment). All mortality that is not derived specifically from fishing is included in natural mortality. This will be included in what we might want to do from here and the group will determine if these test adequately represent these additional perturbations. Butterworth returned to Figure 14.

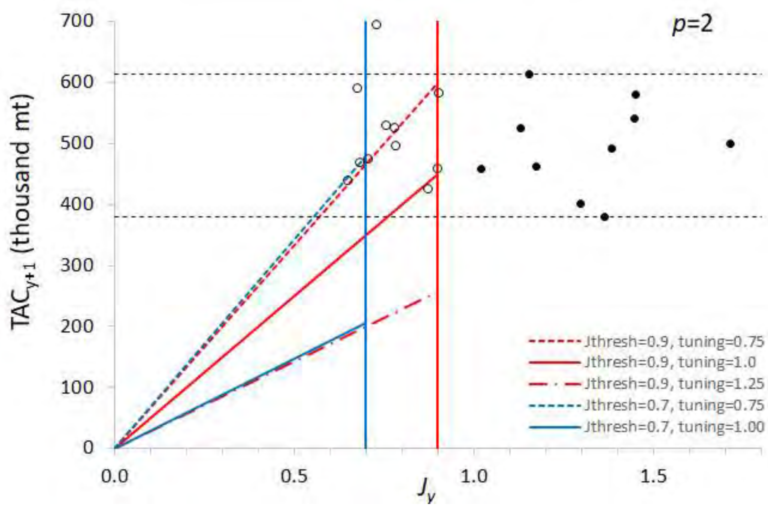


Figure 14. Illustration of the MP variants' rule for two value of the $J_{threshold}$ control parameter, with the γ control parameter value tuned so that the median lowest SSB for Robustness test 4.1 (t) is equal to 0.75, 1.0 and 1.25. In some cases, the target could not be reached ($J_{threshold}$=0.7, t=1.25 for p=2 and $J_{threshold}$ =0.7, t=1.00 and 1.25 for p=3). The horizontal dash lines show the 2009-2018 minimum and maximum landing values. The last 10 years' historical (2008-2017) J_y vs TAC_{y+1} are shown as black dots (Figure 1 from the *Further resultsA_18 July; Appendix B*).

Hanson observed that overall, starting with premise of this exercise of when to put in control rule or TAC to maintain biomass, so far the robustness analysis is good. If we are going to keep to 400-600K mt to maintain the biomass, a HCR like this looks like it would work well. If **Hanson** was to develop his own approach it would be to determine what level of SSB or barometer of biomass we want to maintain both at the lowest end or start at the top end and then set a catch rule to maintain that biomass where it is. He would have looked at a similar approach but from the other direction and put 'guard rails' on the fishery to be sure they do not exceed a mortality that would even trigger something. **Hanson** is still not clear how what he is viewing will be used to gauge if we are overfished or overfishing as these terms are used in the assessment. That needs to be where the cutoffs are for the SEDAR process to know what the status is of the population. **Butterworth** suggested that this is more of a philosophical debate, and that the MSE approach is coming from the other direction. In other words, the MSE informs you of what should be defined as overfished and overfishing, according to your control rule, rather than defining overfished and overfished *a priori* and then designing a control rule that avoids these conditions. But, the approaches are in fact similar. The approach **Hanson** is referring to takes the assessed stock level and then sets the catch. You could do that here. In the same Figure 14, ignore the threshold and replace the J_y index with SSB and you would get something similar. You could change the straight diagonal line in this figure and change it to go to zero at something like SSB at 0.3 rather than through the origin as it is now. A simulation approach (i.e., MSE) to test the performance of those measures would likely result in the same scope of catches.

Jones further addressed **Hanson's** point. The first concern is to keep population in a desirable range (not overfished). The second is that the stock status is informed by the assessment, which determines whether it has fallen below some SSB level, in which case the stock would be determined as 'overfished'. When we set up a HCR like this, the population status is somewhat analogous to the index approach. **Hanson** wants to maintain biomass at some sustainable level rather than evaluate the effects of other parameters ON the biomass. **Butterworth** stated that a primary issue with small pelagic stocks is whether you can even determine MSY, but if you could, you could set a target level of SSB at MSY and see if the projections beyond the terminal year in the assessment results are able to maintain those levels in the Base Case (Figure 15). The top line in the panel is SSB out of the benchmark assessment. A horizontal MSY line could be laid over this graph to see if the projections are maintained above or below, but again, MSY is difficult to determine in Menhaden. B_{MSY} is heavily dependent on the models being used.

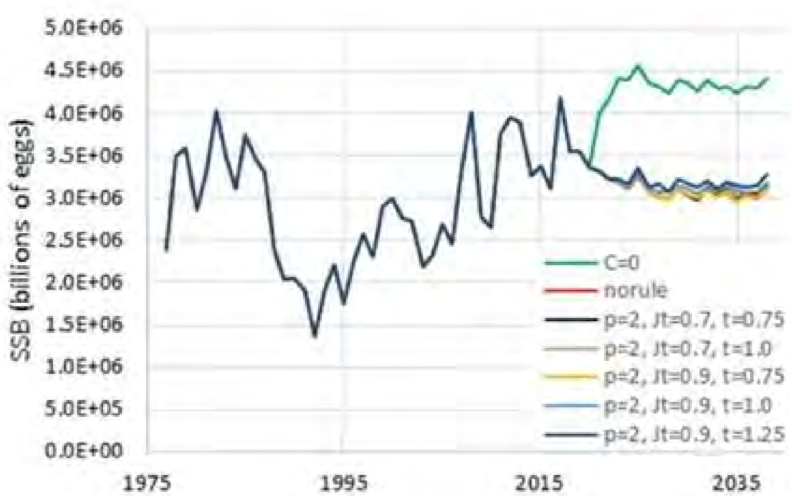


Figure 15. Historical estimates and projected 20-year for a series of quantities for the **Base Case** OM and **Robustness tests, 1.5 and 4.1**, with no future catch (green lines), without (red lines) and with a series of the MP variants with p=2 (Figure 4a upper right panel top row from the *Further resultsA_18 July; Appendix B*).

Jones felt that the answers **Hanson** is looking for are actually in this analysis. Maybe the data are available, but the equation should be different per **Hanson**. All the scenarios that include fishing drive it down so it still begs the question of what level of biomass should we maintain in the water and then restrict the harvest based on that. **Jones** said that this was done using the best information available and, in every case, there is no evidence that the HCR exceeds an acceptable range. **Butterworth** said that these objectives have been implicit in what we have been saying. In the figure above (Figure 15), if we stopped fishing completely, the fishery would be pristine with no catch based on the green line which would put SSB at around 4.3 or 4.4 x 10^6 billion eggs. The cluster of lines using the HCR result in SSBs around 3.1 billion. If we use an F40% target, which are widely accepted and used in many fisheries management situations, we would be WAY above that target, we are more like 70%. You need to compare the no fishing level to the HCR levels. This ratio will determine how well we are doing relative to what is possible; we are around 60-70% in most scenarios. Even for forage fish and MSC standards which would include the ecosystem targets this seems pretty reasonable. The MSC interpretation of our results will likely all be based on the Base Case. We are forcing in additional uncertainties that are not usually included in these considerations.

Butterworth referred back to the basic MP formula in Appendix B in the July paper. This is just the rule for setting the catch.

If $J_y < J_{threshold}$:

$$TAC_{y+1} = \gamma J_y \qquad \text{(B.1)}$$

where

TAC_y is the catch limit that applies for year *y*,

$J_{threshold}$ (no units) and γ (units: thousand mt) are control parameter (tuning) values (the initial choices (baseline MP) are $J_{threshold} = 0.8$ and $\gamma = 500$); Figure 2 illustrates the rule for these choices for these control parameter values, and

J_y is a measure of the immediate past level in the abundance indices that are available to use for calculations for year *y*:

$$J_y = \frac{1}{p}\sum_{y'=y-p+1}^{y}\left[\left(w_{gill}\frac{I_{y'}^{gill}}{I_{2017}^{gill}} + w_{seine}\frac{I_{y'}^{seine}}{I_{2017}^{seine}}\right)\Big/\left(w_{gill} + w_{seine}\right)\right] \qquad \text{(B.2)}$$

with

I_y^{gill} and I_y^{seine} being the observed gill net and seine indices, respectively, in year *y*,

He stated that you could come up with other rules but explained the formula again. Whether you use the assessed biomass or the indices, whether it is a straight line or a hyperbola, you end up with is a plot like Figure 16. No matter what the HCR, you do not want either the catch or SSB to drop too low. You will end up with charts like these regardless of the formula or index you use; any acceptable rule they will all give results like this. The consequences of any simulation will look like this and with the SSB line at 4.2 with no fishing still puts us around 65% under this MP and HCR. These charts are giving the answers to **Hanson's** questions according to **Butterworth**.

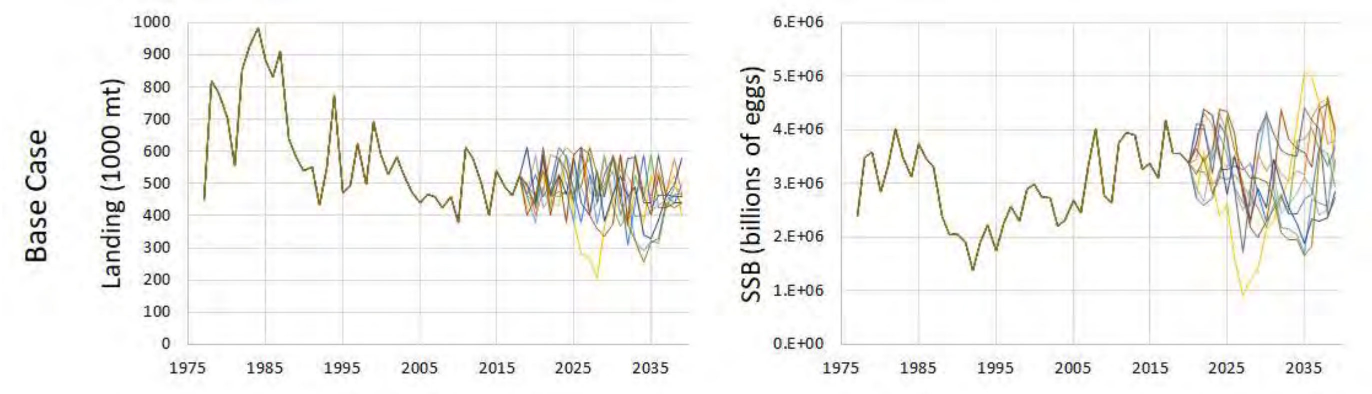


Figure 16. Worm plots for projected landings and SSB for the advocated MP variant with *p*=2 for $J_{threshold} = 0.9$, for the Base Case and Robustness tests 1.5 and 4.1 (Figure 7 top row from the *July paper_fin; Appendix A*).

Jones noted that for the Robustness tests, even the unfished biomass levels also go down. The questions **Hanson** is asking can be examined with the tests we are providing.

Hanson asked if, when the abundance is below that index, where would you set the catch? How is F coming into play here? **Butterworth** revisited the formula above and noted that if we change the index to biomass, gamma would be fishing mortality. Same formula. That is why at its root, it is a constant fishing mortality model. Just to be sure you are comfortable with this, Figure 15, if you replace the Base Case plots with the ratio of $SSB/SSB_{without\ fishing}$ which would put you back at around 65-70%. You can look at every line in the projections for the Base Case and the Robustness tests at the percentage you end up compared to the 'no fishing' line. In the Robustness tests for 1.4 and 4.1, you must judge the percentage based on the no fishing line (green) in the test panels, not the original Base Case. SSB is going down but do the HCR options still maintain the acceptable percentages? That is what you need to look at.

Jones provided some additional thoughts on the F question usage. The addition of a hyperbola/curve for the MP rather than the diagonal through the origin would no longer be a constant F approach. **Hanson** is thinking about a fluctuating F as it relates to the biomass and the catch limit associated with the F so that the F is proportional to the biomass and the catch will be set based on that. **Butterworth**, regardless if the horizontal axis is the index or biomass, it is the same. **Butterworth** suggested that **Hanson** might like to evaluate a rule that ends up as a parabola which **Butterworth** and **Rademeyer** produced but was not considered in this final analysis. It would make fishing mortality go down as biomass goes down. It was published in another paper and **Butterworth** would be able to provide the document via email if anyone wanted it. **Mareska** reminded that the fishery is conducted with spotter planes so that even if biomass goes down, the effort will not necessarily go down. That is more a CPUE issue, not F however. **Butterworth** will include **Hanson's** request for more performance statistics similar to what **Runzel** would like to see in additional robustness tests as the team moves forward after this workshop. **Hanson** also noted that the HCR does not kick in until the next year after the index is already on the downward trend. Looking at the wormplots in Figure 16, what would be a TAC to prevent the dips that we see in the plots that also consider the other parameters besides fishing that are affecting biomass. Why not have a HCR that is always turned on, setting at TAC at 400 or 500K mt that would prevent even getting to a point where those dips might occur and minimize that step decline.

Butterworth reminded that we do not really know what is going on, all we know is the indices and try different versions of rules but the future is still unknown with a lot of uncertainty. If we look at the very first plot in Figure 14 (above), depending on what the status of the resource is, we are responding to that status with more of less risk based on the tuning of the diagonals. **Jones** noted that **Hanson** would like something that is turned on earlier so there was a limit already in place to protect the resource from those times when there may be a downturn. **Hanson** would like a level that the industry can keep fishing for more years but at a lower level or quota. It is more of an accountability measure to prevent getting into a lower quota situation. **Jones** pointed out that the declines in the robustness tests had nothing to do with the quota or fishing, they would happen without fishing. The HCR just prevents it from declining even further by adjusting fishing and encouraging recovery.

Adriance pointed out that if these types of things were happening (e.g., increasing M or persistent low recruitment), assuming they are even plausible, they would be noticed at the state and other levels and will be addressed quickly by management anyway. He thinks we are getting wrapped up in the extreme cases. **Butterworth** said consider the fact that we only know what we know. An assessment provides a best estimate, not exact status. If you get a few surveys where recruitment appears low but it is fine, in other times, you may have very bad recruitment but the surveys are still catching some larval or juvenile animals that suggest otherwise. We need to be able to consider the signal to noise ratios we get out of the data. That is what is driving this.

Cufone had invited two of her students, **Alyssa Conti** and **Jennifer Stucker** from Loyola Law, to sit in on the discussions in the afternoon. They were introduced along with the other participants.

Break

Jones would like to revisit other uncertain future scenarios that are not included in the robustness scenarios already presented here. Are there any other management issues missing that we need to discuss or perhaps include in additional results? Also whatever management strategy gets put in place, something could happen that is outside the range of conditions used to identify the policy. If we find a set of reasonable policies, how would the model be used to effect these policies? What role would the agencies play or the industry play? Finally, are there any additional performance measures other than what the analysts have already provided?

Jones re-displayed Figure 17 which shows what these management procedures mean. In any given year, fishery-independent data is collected in surveys run by the states to derive the index which is J_y. The index can be averaged over two or three years so that the decision is made the following year about if and how much the harvest might be restricted. If the average was 0.6 (green) and the solid red line is the policy in effect, then the TAC (green dashed) for the coming year would be around 300K mt. If the following year, the index is 1.1 (orange), it is above the red line which means there would be no TAC for the subsequent year and the harvest is assumed to lie within the range from 380 to 610K mt based on recent history. Each year the index would be recalculated and the catch would either be limited or not limited depending on that index value.

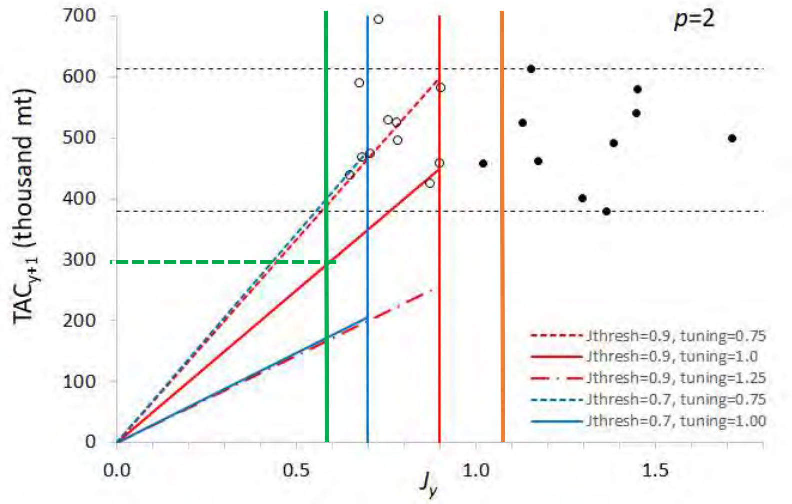


Figure 17. Illustration of the MP variants' rule for two value of the $J_{threshold}$ control parameter, with the γ control parameter value tuned so that the median lowest SSB for Robustness test 4.1 (t) is equal to 0.75, 1.0 and 1.25. In some cases, the target could not be reached ($J_{threshold}$=0.7, t=1.25 for p=2 and $J_{threshold}$=0.7, t=1.00 and 1.25 for p=3). The horizontal dash lines show the 2009-2018 minimum and maximum landing values. The last 10 years' historical (2008-2017) J_y vs TAC_{y+1} are shown as black dots. (Green and orange lines added by VanderKooy for discussion above only)(Figure 1 from the *Further resultsA_18 July; Appendix B*).

Hanson asked where the 0.9 threshold came from; was it intended to bracket along with 0.7 around the original 0.8? **Jones** confirmed that the modeling group chose these. **Butterworth** explained that going greater than 0.9 had no resource or conservation benefits but between 0.7 and 0.9 there were effective tradeoffs between the objectives for the resource and the industry. Anything lower than 0.7 has detrimental effects on the resource by not acting early enough to provide any recovery. **Dix** asked about no additional conservation benefit. Does that mean that at a certain point of reducing fishing pressure there is not a response? **Butterworth** explained that if you put in a threshold and set it higher, you do not have to reduce catch much, but it just provides a tiny restriction more often, so there are greater management costs. **Moncrief** pointed out that everyone needs to have the same background as these originated from as far back as the assessment. If you look at the SSB projections panel (Figure 15), it shows that since 1993, there's been an increasing trend in SSB over that time period. Looking again at the historic combined index (Figure 18), we have been on average, between 0.7 to 1.7 while the landings have been between the values in the Figure 1 at 390 to 610K mt and, in fact, the SSB has been increasing over that time. We are here to figure out a HCR that will continue the process as it is and prevent reductions in SSB should something in the environment or recruitment change. There has been no catch limit on this fishery to date and it has not been necessary thus far.

Figure 18. Historical combined index J_y values for p=2 and p=3 (Figure 2 from the *Further resultsA_18 July; Appendix B*).

Runzel: "Once we get an index above 0.9 under this HCR, there is no cap or restriction applied to the fishery because it is not necessary?" **Jones** confirmed and stated that was the point **Moncrieff** was making as well. There has not been a need to implement any restrictions in recent years and SSB continued to increase over historic levels. **Butterworth** pointed out that these rules are put in place assuming the fishery will continue to operate as it has and nothing changes significantly in the environment. **Mambretti** is concerned about which years are used for the averaging in *p* values of 2 or 3 because including THIS year's young fish will not contribute to the SSB until the following year. Perhaps the index of recruitment needs to be weighted differently. **Mambretti** feels that the current and previous years are probably the most important year classes for next year. **Butterworth** reminded that this year's fish are being applied to next year. If the index is low, it will mean next year there may be less fish so you take less this year to ensure you have some additional spawning adults out there next year.

Jones reviewed Figure 13 (above) representing impacts of the various MPs in the Base Case and two Robustness tests for those who were not here to see it earlier today and opened the floor for potential questions or additional comments. **Butterworth** indicated that in most of these panels, the policy and tuning you would ideally like would put a dot in the upper right hand of each plot to ensure the best outcome for both the resource and the fishery. However, that is not likely, instead you run along a negatively sloping diagonal line. **Jones** admitted that looking at this for the first time is probably a little overwhelming but depending on your general values, you would likely fall more to one end of that spectrum or the other; allow a little more risk on the harvest versus a little more risk on the resource. Are we able to find something "in the middle" that represents a policy we can live with? **Runzel** indicated the lowest biomass in the history of the current time series was 1.5 and the median final SSB projections panel for the Robustness test 1.5 ends up below that historic low under the mortality increase case. **Jones** stated that these ARE the most extreme cases we are using so we still need to consider the plausibility in these examples as well. **Moncrief** noted that it might be worthwhile to go back through the plausibility of these two example scenarios.

Jones reminded that 1.5 was a 20% increase in natural mortality annually for 20 years of projections. The consequence is that unfished SSB would be half as great by the end of the series, relative to the present. **Leaf** reported that there is a very diverse field of predators from the 50 years of diet studies that have been done and Menhaden is just one of the prey reported. If there were increases in all the predator populations simultaneously, you might see something similar but in most ecosystems, not ALL competitors can increase, there are some that would decline. The productivity of the ecosystem would also have to decline at the same time. There are actually multiple factors that have to occur together for this extreme case. **Hanson** wondered if we would gain any insight looking at spikes in mortality, which might be a more plausible scenario. Likewise, punctuations in increased mortality at the scale of a few years of slow decline might be more plausible rather than five consecutive but a large decline in a single year or two. Instead of recruitment declining by half each year for five, we would have a single large reduction.

Butterworth returned to Figure 6 (above) of the hockey-stick graph. You could find a few years of bad recruitment in that scatter which is worse than a few you see in the other plot. **Hanson** is looking at a situation where there was a true recruitment failure in a few years or a couple of years.

Dix stated that ten years ago, very few people would have said we will have a series of the hottest years on record or that the Bonne Carrie Spillway would be opened two years in a row and twice in one year. At this point, any of these scenarios is plausible. You cannot say something bad is not going to happen. **Runzel** stated that environmental perturbations are probably much more likely than sudden increase in predation at this point but it is better to plan for future than to be stuck in a scenario where population ends up dropping. **VanderKooy** reminded that the index is being determined every year and the response is not static. In the event a trend in the index began to occur, there would be a reevaluation of the MP and a new stock assessment to revise the conditions we are representing in the model.

Dix said we are talking about robustness tests that the analysts came up with including scenario 1.5 and 4.1. These are not necessarily dooms-day stuff. **Hanson** addressed **VanderKooy's** comment and suggested that with the expectation of multiple assessments over the next 20 years, we should probably consider a punctuation event in mortality or recruitment. **VanderKooy** stated to **Dix**'s point that the two extremes which are here are plausible but that we would see these changes sooner than five or ten years from now. We would stick to the MP that was agreed to but we could readdress the management objectives if necessary. **Jones** agreed and stated that there is a level of comfort with that in that the index is a way to see if and when a problem starts to surface. **Moncrief** pointed out that the process is dynamic and we will always be reassessing it. Also we need to look at the control rule as well as the scenarios and how it would help. **Hanson** asked what about periods of prolonged catch as a Robustness test – What happens when there are multiple years of a catch 600K mt or above? Does that affect the biomass. **Butterworth** reminded that we have not actually come to exceptional circumstance yet but the assumption in this model is that we remain within the range of conditions observed in the past. If we exceed the range for harvest beyond the historic amount, we likely would need to determine if the model still represents the population. **Dix** is concerned that the projections may not actually be serious based on what he's hearing. These might change if the strategy changes? **Jones** cleared it up that if there is a change, the managers would begin to process the new information and adjust the MP accordingly if it deviated from the historic norm. **Runzel** asked if we are going to develop a process for another evaluation. It keeps getting mentioned but does not appear to be developed yet. **Jones** and **Butterworth** indicated that there would be a formal process should exceptional circumstance occur.

Discussion of How to Handle Exceptional Circumstances

Butterworth returned to the July paper and began the overview of the exceptional circumstances in Figure 19. The critical point is that we have put something in place including uncertainties and projected based on those which were derived from the assessment. We are also able to predict how much we exceed the limits above or below for the two indices going into the combined J_y. The whole point is not to have an assessment or an exercise every year so that time can be spent on research. If you are borderline on any of these upper and lower limits it means you start to look at the MP and potentially review the MP and HCR. You do not want to have to review the MP every time you have something outside the 'norm' but watch and then if it persists, make the re-evaluation. Rules for when and how to re-evaluate using a cogent argument have been developed into an approved methodology for moving forward for other fisheries. These protocols are routinely copied from the Australian approach as a template.

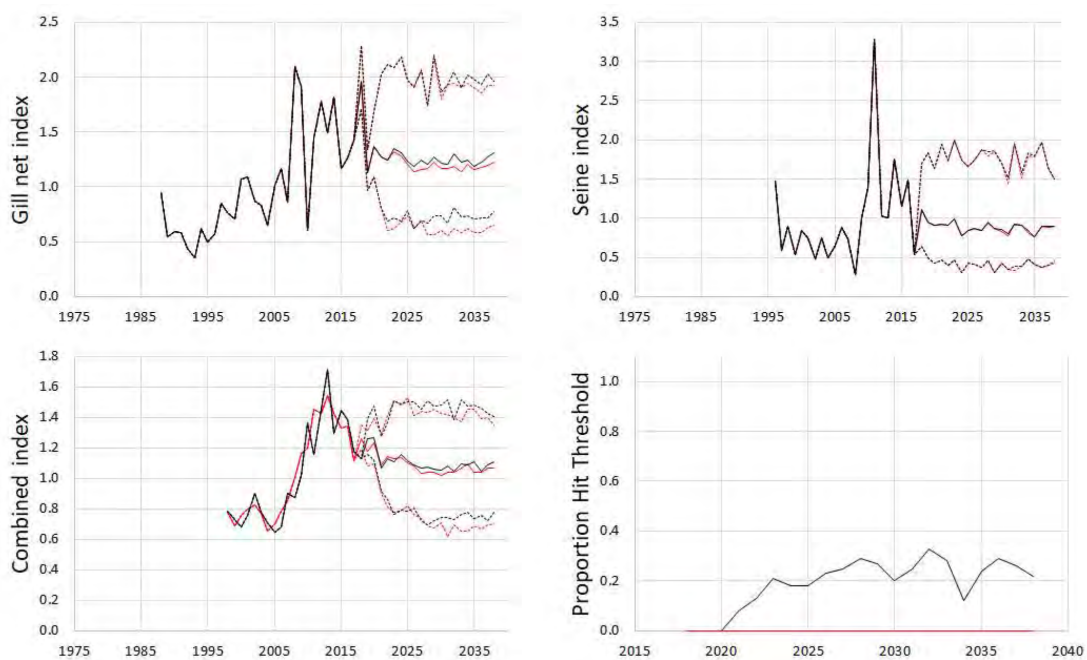


Figure 19. Historical estimates and projected 20-year median with 10%- and 90%-iles for a series of quantities for the Base Case OM, without (red lines) and with (black lines) the advocated management rule (Figure 8a from the *July paper_fin; Appendix A*).

Existing MPs are routinely reviewed but the time between reviewers can vary from four to nine years. **Jones** asked if **Butterworth** could provide some examples of exceptional circumstances. **Schueller** stated that there are alternate benchmark and update assessment every three years approximately, so a process is already in place that could be dovetailed into a MP review. **Hanson** asked **Schueller** if benchmarks may be less often and updates more often like the Science Center is doing with their research track assessments. **Schueller** said that she is the only resource available for Atlantic and Gulf, so no, not likely more often than currently but the SEDAR terminology is not necessarily applicable to how we are proceeding. **VanderKooy** explained how the benchmarks assessments are currently paid for. The Commission covers the Data and Assessment Workshops, and NOAA/SEDAR covers the cost of the CIE review and the Review Workshop. That is not going to change but updates are conducting through the Commission directly and entirely.

Hanson wondered if exceeding the 600K mt upper harvest level would be considered an exceptional circumstance and if so, perhaps we should just make that the cap and stay ahead of future impacts due to increased harvest. **Kuttel** argued that there is reason the catch stays in the recent historical range (capacity) and an arbitrary cap should not be done just to make us feel good. If the index indicates that the population is good, why restrict just for the sake of restricting, you might as well cap it 400K mt then, since it is completely arbitrary. **Jones**: "Would not you need some reason to believe that higher landings would be somehow deleterious to the environment in some way?" **Butterworth** cautioned to be careful saying exceptional circumstances. Not everything is exceptional but you would potentially look at triggering some sort of turn-the-crank assessment and only if the SSB was declining, then you would maybe reconsider the MP. But if high harvest was understandable because recruitment continued to be high or increasing, take that as a benefit for the fishery rather than something negative. Common sense has to be applied to the trends and status. **Dix** asked if there is a way to model on past events so that if we looked at what affect a cap would have had in the 1980s and see how the population would have projected out and perhaps it could have prevented the stock from 'bottoming out' in 1992? **Butterworth** stated yes, at a formal level we could look retrospectively. This has been done on European fisheries by **Butterworth** and colleagues about seven years ago. Our Base Case actually does that already, assuming the historical parameters remain the same but the reason we want the rule is that the future might be different.

Jones asked if the modeling team had thought of doing a robustness scenario where the range of landings was larger than those bracketed over the last 10 years? It seems like you could do a process where you are selecting from a range that is higher than 610K mt… maybe 700K mt. **Hanson** stated this would be helpful at addressing his concern about the impact of increased harvest. **Leaf** pointed out that increased natural mortality could be considered analogous. In that test, the fish are dying at the hands of increasing predators but the fishery is already a predator of sorts and you could infer that into that scenario under total mortality. **Butterworth** said there is also a test in place already for undeclared catch which increased fishing mortality. If this exercise may make those who are concerned happier, it could potentially satisfy the group but could hurt the process by including something that the analysts have pre-determined it has no additional value. **Jones** said that perhaps the team should explore that. **Schueller** stated that for the last few years, there have been only three plants in operation and they have a limited number of boats. It is possible that there could be more plants in the future but there is an upper limit on just how much catch could actually be processed at the plausible capacity. The number of boats should also be taken into consideration. **Himchak** stated that setting an upper bound on the harvest would be triggered by the HCR. If you were harvesting over the upper limit the MP has considered, how would that be determined as doing harm? **Jones** acknowledged that occasionally he has run a model and been surprised by the outcome so just because we believe things will respond in a certain way, it does not always happen. Sometimes it is worth pursuing if even just to confirm the analyst's suspicions. **Dix** asked **Schueller** what the upper limits of plant processing is based on – is it capacity or economics? **Schueller** said the plants can just push through just so much in a day so capacity is limited. If multiple boats arrive at the dock, there is a limit to how much they can pump and process. We do not know what the number is. **Dix**: "This goes to the point raised by **Cufone**, if demand goes up we could expect more plants?" **Landry** stated that more production plants are highly unlikely due to the expense and Coast Guard regulations. The choke hold is the plant not the vessels or availability of fish but rather how much you can push through. **Cufone** stated that if you are already running plants at full capacity, could not it be capped at the choke point level of catch? **Landry** stated that the harvest cap should be adjusted at the resource not based on the limit of the industry. **Cufone** asked what the industry felt the cap should be; higher or lower than what was already suggested? Is there a better way to determine what the upper should be rather than just keep talking circles? **Kuttel** stated maybe we are getting confused, we have a model to set reference points based on where the fishery has been over the time period and there is no damage done. There is not enough data to state that we should go no higher. We need to get to a point where we have a demonstrable cap, we just do not have any data that provides that now. **Mareska** asked that if we could run the analysis to look at what the choke point is for the fishery and it turns out that the upper limit is actually much higher, maybe up around 900K mt, would the ENGOs accept that as a potential cap? **Cufone** said maybe but that we need to find another way because we are getting nowhere in determining a cap with the current modeling. We have spent two days hearing the environmental situations into the model and that there are issues with the how the model could or could not handle those, we need to find a way to actually include those. We need to get to some smart, forward thinking management. **Adriance** reminded everyone that the model is only intended to test a HCR, not a cap. Setting an artificial cap is not the intention in this exercise. **Kuttel** said that, even if we went over 600K mt and it has a negative effect, this will be shown in the indices and the HCR would kick in. That is the intention of the rule is to respond. **Cufone** wondered that if that is the case, how long does it take for management to react in the event and how long does it take to manage the change? If we cannot figure out a way to pick a rational set of numbers for management, we need to do something else. **Jones** stated that we are trying to figure out a way and are moving in that direction. **Butterworth** stated that everyone is going a little beyond the intention of the rule. The index and the rule only apply if things continue on as they have so going over 600K mt might show up in the index but we would need to review the model and be sure it can respond outside the historic parameters we have set. We first must just agree on what to run and run them. We do need to be reasonable with how many additional things we explore and be rational in the value of those parameters and scenarios. To **Cufone's** point, this MP is a process that is potentially triggered annually and if management does not respond quickly, the MSC will. They audit the certification and the ability to respond to changes is part of their requirements for keeping certification. The industry has a serious investment and wants an effective rule and response with regards to any MSC certification and they could lose it if these guidelines are not met. Butterworth suggests letting the catch increase a year and see what actually happens. **Dix** thinks we are creating a prophylactic for the fishery to protect them but it seems like an arbitrary cutoff to only include the last ten years because we HAVE gone over 600K mt in the past. You can say it is arbitrary but it is more prophylactic and coming from a place of bad faith. **Jones** asked if we used the models to explore further the higher catches of the longer past, would others be satisfied with that? **Hanson** and **Cufone** agreed that they would. **Cufone** pointed out that we are super slow in management in the US and the MSC provides a lot of warnings but cannot be the manager. The agencies, NOAA, and the Commission need to be the management bodies. Changes cannot take many years to implement. **Himchak** pointed out that MSC does provide supervision. They have a third party assessment team that reviews and provides the recommendations TO the MSC.

Butterworth will give **VanderKooy** an example of a paper that is the MP they used in other fisheries as homework for Friday morning (The 2018 Operational Management Procedure for the South

African *Merluccius paradoxus* and *M. capensis* Resources – referred to as the *2018 MP paper* henceforth). It shows how the MPs look in practice.

Jones wrapped it up and pointed out that it is clear that we will not come out of this workshop with some consensus on a potential HCR but that we will have hopefully clear direction on additional tests that need to be explored to get to a MP that makes sense and gets us to a decision.

Public Comment

There were no comments received as the public was at the table and there were no participants on the webinar portion.

Adjourn 5:00 pm

<u>Friday, July 19 (8:30 am – noon)</u>

Jones summarized where we have been to this point in the workshop and the necessary steps we need to complete to get to the finish line. We have had the operating model presented to evaluate the performances of potential procedures and HCRs looking at the best guess for the future and a couple of scenarios which could impact the resource in the future. We discussed the tradeoffs between meeting the objectives of the resource and the opportunity to extract economic benefit for the industry. What we need to spend a little more time discussing further would be what considerations we evaluate as exceptional circumstances. A wise rule has to include the assumption that things could be very different in the future than we expect. We need to discuss and agree to any additional simulations we would like the analysts to consider as tests of robustness. Finally we need to discuss how we move forward to implementing should we come to agreement on a MP or HCR.

Himchak noted that fishery-independent data is at the core of our current MP but there are issues with the surveys in the long-term. Under exceptional circumstances, do the states foresee any situations that could undermine the index itself? Can it become less informative moving into the future? **Jones** wants to be sure we include correct items as exceptional circumstances but not the integrity of the data collection itself. What conditions could be seen IN the surveys that would potentially cause changes in the confidence of the index or the ability of the index to continue to represent a correct state of the population?

Butterworth stated that the index or survey comparability would always include noise annually but at what point does the noise indicate a problem. We probably cannot do much more today but need to include that in the list of topics of exceptional circumstances. We also need to include it in review discussions in the future.

Dix asked about the modeling team running tasks beyond robustness. He would like to include modeling of the population changes in the older history. The recovery for Menhaden took place in the 1990s after much overfishing when we had very low predator rates and now we have recovery in other fisheries that could be impacting Menhaden and this needs to be better understood. Also micro plastics have been found inside Menhaden. **Jones** said that our modeling approach is not going to address some of **Dix** questions, this would be more in the realm of ecosystem modeling like what **Chagaris** is working on. This effort is very much focused on a single species model, just as is the case for the benchmark assessment so the approach is not suited to addressing ecosystem-scale questions, at least directly. Work by others will get to this but not using this approach. **Dix** would like to see the HCRs run on past circumstances if that can be done, to see if rule would have kicked in and how it might have affected recovery or prevented the collapse in the early 1990s.

Leaf pointed to the *Ross-Gillespie et al. 2018; Appendix D* paper provided overnight by **Butterworth** which puts some boundaries on exceptional circumstances and would help us in this discussion. We need to use these terms and conditions to guide where we go with this discussion. **Hanson** stated that the model can be used to encapsulate these things that **Dix** noted and perhaps modify the level of risk as we see more uncertainties come up and are not directly modeled. **Butterworth** stated that what has been raised is that there could have been different levels of mortality in the past. We have just been looking toward the future. If natural mortality had been different, we may see different trends. **Jones** stated that the modeling group did some robustness tests for changes to natural mortality but it did not meet what **Dix** was talking about necessarily. One could do that type of test by surmising that the mortality rates were lower in the 1990s and higher today. **Butterworth** agreed, and noted you could get into this process with a much more complex multispecies model or allow for a change in natural mortality could be proxy for changes in the predator community.

Butterworth provided an overview of the *Ross-Gillespie et al. 2018; Appendix D* paper which was used for species off South Africa and which explains how the MP was applied. It details how the TAC was calculated, how the index was derived, and what to do with missing data. It is important to have a plan to deal with a missing survey or certain data points. Most of the document is appendices. Appendices A, B, and C give a lot of detail as to how to go from raw data to input data which is what **Himchak** was concerned about. These lay out how to standardize WHAT to do in the MP and in the event of a change in the uncertainty. How are the various surveys used and applied in the MP under the current form. This needs to be consistent.

Appendix D points to how to review the model when there is deviation in the MP output. This is a process document borrowed from the Australians that says the rules cannot be 'tinkered' with but requires the use of common sense and scientific judgement. The word here to note is 'compelling', such as survey results that are outside the norm. The details simply layout HOW to proceed with re-evaluating the MP. This is the sort of clearly laid out plan we need to develop and agree to for working with our MP and HCR in the future, so everyone knows what to expect when circumstances change. Finally, Appendix E explains how to project and interpret the forecasted data. On Figure 19, the red dot is what you fill in each year and see if you remain within the colored range of limits. If you get outside the limits, you need to determine why and consider reviewing the data and MP to determine if something has changed so the projections are no longer representative.

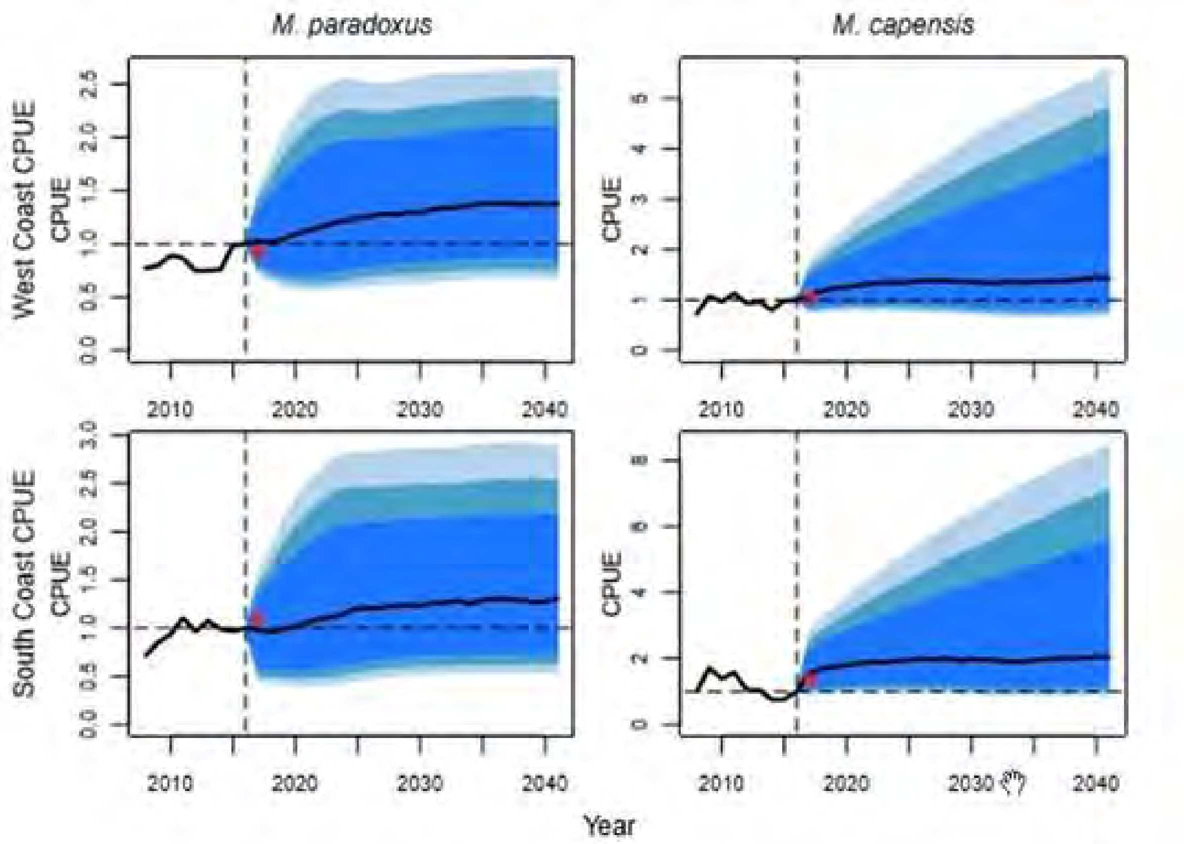


Figure 19. The 95, 90, 80% PE and median for the projected GLM-standardized CPUE for *M. paradoxus* and *M. capensis* for the RS under OMP-2018. The red dots show the 2017 CPUE indices, standardized relative to the 2016 value in the updated GLM series (Fig. E1 from *Ross-Gillespie et al. 2018; Appendix D*).

Jones returned to ask for any additional thoughts on the criteria laid out in the appendices that might be something we need to track. What things other than index and catch do we need to track moving forward to determine if we have entered into an exceptional circumstance? **Hanson** asked if the seine index can be used as recruitment index on its own between assessments. **Jones** stated that is currently being used as index of recruitment. **Hanson** asked how we would do this? He suggested that maybe we need a standing technical committee like the Council's SSC and maybe suggest changes that should occur to the MAC so actions could take place. **VanderKooy** answered that historically, this was what the Commission's Technical Coordinating Committee (TCC) did. Since we do not manage actual species, the TCC reviewed the available data and developed data collection programs. We are in the process of regrouping today, looking at what their role will be in the future. At the MAC meetings, we used to get a forecast from both NOAA and the LDWF which would provide the agencies and industry some idea of what they might expect in the coming year. It is unclear however who would take the data and put the index together on an annual basis to present to the MAC during their March meeting. If the MAC found a problem, the TCC could implement a review of the MP or recommend additional data gathering.

Dix asked if this modeling is going to be peer reviewed. **Jones** indicated that the process, not necessarily this model, has been reviewed. **Leaf** answered that was part of what the technical team did in the process. **Dix** suggested that someone else outside the team might be needed to review. **Butterworth** noted that an outside review would be helpful but the trouble is finding someone independent who understands the biology and model to do a sufficient job. You want the review but having outside participants IN the meetings and model development is more valuable. Sometimes the feedback is helpful but best when all participants are in the same meeting. **Jones**: "No one disputes the merits of an independent review, but that is probably in the next steps." A third party oversight might be useful to consider. SEDAR uses the CIE and maybe there is something analogous that could be developed but participation IN the meeting is critical in understanding how the MP was derived.

Hanson pointed out that timing would be important in relation to the benchmark and assessment processes but it seems like this process would be part of that so the reference points can be used to gauge the status of the stocks. It would be good to get this implemented prior to another benchmark SEDAR review. **Leaf** stated that these logistics are a challenge given time and expertise necessary. Could the MSC review and process be satisfactory? The **ENGOs** indicated a strong no. **Landry** stated that an independent group from Dublin was employed here for MSC and that should suffice. **Hanson** - that is not the role of the MSC. They are not judging the stock assessment. **Dix** compared the MSC review to the investment confidence ratings going into the financial collapse. **Jones** summarized that the MSC review and the technical process being done here are two different things. **Butterworth** said that the MSC process does not have time and people to do the degree of thoroughness and detail. They are looking at an overview and would pick up on bad oversights.

Jones tried to move the discussion forward towards any other important criteria for exceptional circumstance. **Runzel** said that predator impacts are not well understood and if ecosystem conditions change in the future and we do get more information, they should be addressed. Species like brown pelicans are pretty key as predators. **Jones** generalized in his criteria list interpreting **Runzel's** concern. **Leaf** pointed out that things like predation will come through in an index of abundance. **Moncrief** asked if there is any detailed information about the diet of brown pelicans and finfish. **Leaf** stated that at some parts of their life cycle, Menhaden are fairly important to brown pelicans. **Leaf** stated yes, we have limited data but we do not have a reliable time series regarding brown pelican abundance further back than three or four years ago. **Runzel** stated it is in Juliette Lamb work and she will provide to **Leaf**. **Jones** understands that the increase of predators in general would show up in increased natural mortality as a proxy since that is where the predation would be seen. **Dix** asked whose responsibility it is to gather the lacking data. Is there a plan to get that data and whose role and responsibility is it to get that important diet data? **VanderKooy** stated that these data have come from the academic world – single species studies that were anecdotal and ancillary. Florida FWC has developed a diet lab from fishery dependent and independent data. The FWC has put a lot of work, time, and cost into developing this data in their waters at least. **Jones** wondered if any of this work came out of the BP Disaster to look at diets, feeding, and long-term impacts. **Himchak** stated that on east coast, they rely on the science centers to provide some of this for Atlantic Menhaden. **Hanson** knows that there are restrictions in the funding and these are the issues that need to be highlighted often and pushed from the MAC and the Commission to find ways to fund and collaborate and get diet data. **VanderKooy** and **Leaf** stated that these are high priority needs and have been listed in all the previous Menhaden FMPs and in the assessment results and the first line in all proposals submitted. **Schueller** stated that there have been RESTORE projects submitted but they have been denied. This is a fundamental problem. The PIs, the agencies, and the industry are all helping to get these projects funded by providing info and writing letters of support, etc. **Runzel** would like Audubon to get involved with that process to advocate for getting at these data needs. **Mareska** stated that Alabama got a NFWF project funded to examine stomachs but the actual collection of stomachs had to be piggy-backed off other projects. Mammals, birds, etc. are not being looked at. **Butterworth** suggested another angle, what can you realistically get funded? If it is going to be useful, it needs to be monitored in the long-term, not a quick, single data point. For use in the model, determine those data needs that have real potential for long-term support. **Leaf** pointed out the brown pelican papers highlights that problem exactly. These are not in historical trends and are not useful for our purposes because they are only snapshots.

Jones listed out three topics he carried over from yesterday in moving forward. We have the following:

1. Simulations that consider episodic effects
2. Consider greater range of landings or catches drawn in absence of control rule. At least stretching the upper limit within realistic values.
3. Consider an additional scenario with natural mortality and recruitment in tandem, some interaction of the two.

Break

Jones asked one last time if there were any additional simulations that might be considered. **Jones** had written down the item **Dix** brought up related to simulating past circumstances had the HCR been in place. It was hard to capture but essentially it would be other variations in the current robustness test about recruitment in the future that could be applied to conditions in the past. That would obligate refitting the model as Butterworth had pointed out. **Hanson** is not sure what other combinations of changes beyond just in natural mortality and recruitment. They may not need to be run necessarily but discussed in the context of the results so far of combinations. **Jones** recalled that changes in selectivity patterns were addressed in the robustness tests but maybe that could be in combination should one or more parameters change as well. Consider it and determine if it is worth pursuing at least. **Mambretti** thinks maybe looking at recruitment decreasing like 20% through the course of 20 years projected models with random episodic events included with increased mortality of maybe 1% per year as proxy to simulate a change in the environment.

Implications of Reference Points for Management

Jones followed up that once a single MP is agreed upon, how do we implement? This would probably be more successful if it was done by the industry rather than the states. At the February workshop, the path of less resistance that was noted was to have the industry present to the state agencies and MSC a proposal for a self-imposed MP that the states would consider and judge whether it would be effective and acceptable. Achieving something across all five states legislatively is probably not realistic but self-governance appears to have some merit. We should have a discussion on how that might work and if we can pave the way for something like this to happen.

Hanson would like to see catch limits in place and use the HCR as an accountability index. There needs to be a way to prevent the catch from increasing to begin with. This discussion should be held at MAC meeting to look at data where you get the projections for season and have the ability to set and enforce a TAC before the season starts. Where is the accountability to keep them from ignoring the quota or TAC? The agencies need to be able to go back and adjust the rules on an annual basis to do something like shorten the season or set a temporary quota. This could be done by the agencies each year. **Mambretti** stated that truncating the season a few days would not affect because it is past, monitored by season and on a weekly basis. It would have to be self-enforced by the industry to NOT go over.

Jones would like to hear from the industry and **VanderKooy** as to how this might work, understanding the existence of the MAC and the Commission not having any management authority. **Himchak** stated that states typically do not like to change regulations every year and it takes almost a year to do. There would need to be a MOU among the industry partners as to how this would be monitored. **Kuttel** said that self-regulation of only two companies could easily work together but if another company were to come in, there would need to be some additional hoops. **Landry** stated that industry wants this to be collaborative. There would be a number of setbacks if the HCR is not in place or observed. The industry has an investment and responsibility to help the states figure out that the harvest is being managed responsibly. There are a lot of hammers that would come down if the industry becomes irresponsible. Jones noted that the machinery will still include the public agencies since they are collecting the fishery independent and dependent data and are contributing to tracking the status of the stocks. **Kuttel** stated that industry compliance would be self-evident. There will be a number that will be determined publically and it would be obvious if they did not comply. **Himchak** mentioned the NMFS monitors overall catch also, not the industry. There is independent oversight.

Hanson concurred that the legislative process at the state level is troublesome and clunky but that an MOU precludes the public input and that is an issue. If it is in the state realm, it is much more transparent and without that, it would be an issue. There is no forcing mechanism in place to prevent going over the limit next time and the next. There is incentive through MSC but I do not see the hammer. **Adriance** stated that if there is biological and technical evidence that there is an issue in a Louisiana fishery, they can shut that fishery down within 24 hours. He disagrees that there is no accountability. **Mambretti** stated that since 2008, Texas has had a cap on fishing and every week, the industry has provided their fishing reports to track that cap in real-time. The industry is extremely cooperative and NOAA provides the oversight. The TPWD is a public entity and participates in the MAC and sharing data with the other states. **Jones** summarized that **Hanson** is looking for evidence of accountability whether that is done through a rule process at the state level or through another mechanism. There are a lot of public interests that want compelling evidence of accountability. For this to go forward with public support, there will need to be an accountability mechanism to satisfy the other stakeholders. **Hanson** agrees that there are some mechanisms in place now but there has to be a lot more procedure laid out that explains how all of this works. **Jones** would assume that a draft of an MOU would be worked on and include the information in **Butterworth's** paper with all the appendices.

VanderKooy pointed out from the Commission perspective, we do not have these kinds of mechanisms in place to monitor and enforce quotas or caps. The MAC receives a summary of the previous season and some sort of forecast to determine how the population looks and what to expect after the meeting in March. We no longer have an LDWF forecast like we did from Vince Guillory but the index would effectively provide that similar information. Towards the end of the season, we should be able to get an update on the total landings by the October meeting. That public meeting would provide the groundtruthing to the HCR. We do not have anything formal and we do not have anyone to build that index each year. We cannot volunteer **Schueller**, she works for NOAA so perhaps the states would need to do it individually. We are glad we are finally getting to this point with goals and objectives, reference points for management, and some sort of HCR on behalf of the industry. The MAC reports to the Commission and would have to get permission to cooperate. We need to draft up a process and propose it to the TCC and Commission to move forward using the MAC to provide this oversight. The Commissioners need to be on board. **Jones** remembered that the index from the previous year's survey data could be produced in time for the March MAC meeting. **Schueller** stated that **VanderKooy** was correct and he cannot volunteer her and she cannot volunteer herself either. There is a formal process to get NOAA to agree to have **Schueller** to participate in anything. **VanderKooy** wonders how long would this process take? Could it be almost automated so it gets completed quickly and easily? **Schueller** stated that getting data from the states is problematic but could be overcome. The index would not take long to generate each year. **Hanson** stated that it would be good to sort out the resources. Perhaps an independent survey of three to five years to see if this index is appropriate. We all know that the current surveys are not designed to capture Menhaden. Perhaps a short-term project could be setup to test another survey and gauge the effectiveness of the current index? We are relying on the index as an indicator. It would be good to generate the confidence with some sort of corroborating data that does target Menhaden recruitment. **Schueller** reminded everyone that the seine index is not the only recruitment measure, there is also trawl data and the two surveys were congruent and validated the seine when we went into the previous benchmark assessment. **VanderKooy** noted that this is why the ENGOs need to come to the MAC meetings and the assessment workshops. The states have additional gears they are currently testing but the funding is never certain and the data ultimately must be long-term enough to be used in the assessment. **Hanson** noted that they are calibrating recreational data using short-term sampling and similar efforts would benefit this.

Hanson stated that when this gets into the next SEDAR, we do not seem to have reached a point where the reference points have been set to determine where the stock status can be determined. We need to solidify where are those reference points and how are we using this to maintain the stock above those levels. We do not have reference points here. **Butterworth** explained there are two fishery management camps currently, a traditional camp using reference points and a MP camp which manages to avoid RPs for certain circumstances, using only a limit RP. The performance statistics you predict are the actual measures not some limit and threshold targets. If you prevent negative performance of the resource through a HCR, you do not need to set a limit, the rule does. This is much wider debate in a lot of fisheries right now. **Jones** said that in the US, reference points are defined in stock assessment and the expectation is that the targets will be set to keep fishing at some level relative to those reference points. You could, within this context, set a limit reference point which triggers action. That is basically what we have here with the HCR, an action is triggered when the index goes down. As far as a target, if you evaluate the projections based on the upper range in an unfished state compared to the biomass under the various fishing conditions, its equivalent to an SPR%. **Hanson** said he agrees but we have not codified this next step and this discussion is not concluded. He said that the index is not the reference point for the limit. **Jones** followed that if you look at what they are working on along the Atlantic currently, **Schueller** and the others are looking for ecosystem reference points to determine what the biomass of Menhaden is they are looking to maintain for all their services. **Hanson** does not agree that the index is the reference points for the limits, yes they trigger action but are not a determination of 'overfished'. We need to talk about where that limit line needs to be to get at overfished and overfishing. **Jones** acknowledged that we are trying to avoid getting into an undesirable state with a limit but not making the determination if that state is defined as overfished. That is what **Butterworth** was suggesting earlier. It is in the spirit of the limit reference point. **Hanson** does like the approach, especially between assessments which can be a long period of time. **Butterworth** indicated that limit reference points are problematic elsewhere in the world as well.

Another way it is taken is not the operational point that you should take action but the level at which your action does not want to get below. When looking at pelagics/forage fish, you want to avoid overfishing because of the variability in the environment which contributes to potential poor recruitment. Longer lived species are not as susceptible to these fluctuations. Bottom line is you do not want to drop biomass too low and this MP achieves that.

Next Steps

Jones indicated that these are the tasks for the analysts to address coming out of the discussion so far from these two days. It does not seem too difficult a task. **VanderKooy** will develop a draft report that will capture most of what was said here but it is not a precursor to the modeling team moving forward. Do we see another workshop? **Jones** believes that another meeting would only circle back with more of the same discussions and requests for more analysis. Let's complete the next steps and then formulate a draft management strategy to consider moving forward. We would likely not end up with a 'choice' as a result. We need to rethink what the process looks like once we have completed these additional test. **Butterworth** suggested the analysts, states, and industry come up with something together and not leave the decision open-ended and give a summary with an opportunity to comment. There needs to be a line that stops any additional analyses just to produce more analyses. **Jones** wondered if we want another outside expert to look at what has been done here to evaluate if this is a sound analysis. Would this provide an assurance to the non-technical people more comfortable with the results?

Landry feels like we fine-tuned our understanding during the last five days. Perhaps we can address some of this at next MAC meeting in October. Let's continue to fine-tune and then devote some time to review at MAC meeting. Another multiday workshop would be overkill.

VanderKooy said we need to figure out how to implement this mechanism, monitor it, and enforce it. We need an Appendix D from the *Ross-Gillespie et al. 2018; Appendix D* paper to define exactly how the process will work; to determine where the 'hammers' will be. If it is going to be industry-imposed, will need to be transparent through the MAC so the other stakeholders can witness the process. We cannot have something that occurs behind the scenes which could cast shadows or doubt on the legitimacy of the process. Maybe we can accomplish some of the drafting of this in advance of MAC meeting and whatever we come up with can be presented there. We do not need to wait for the results of the HCR to layout the plan of how to implement it. **Jones** agrees that there is nothing to stand in the way of developing what an MOU might look like while the additional analysis is going on. That would be a pretty good place to be. **Hanson** agreed that this is a good approach. He would like something in writing so they have a draft to review prior to the MAC also. This would be a big step toward achieving transparency. This would be a positive move to include the other stakeholders who may not be able to attend the MAC meeting. **Jones** sees some sort of report from the technical work would result and that a draft MOU or other product would be available for review. We also need to determine how to get someone outside to review the report and the procedures. **Butterworth** noted that if you have only one reviewer, you are stuck with whatever that person provides even if they have not a clue what is happening or are a bit maverick in their opinion. Two might cause disagreement and then you are stuck with a hung review. If you have three, it is costly and timely but more likely to be a better review. The concern is also the cost to getting someone who is qualified. This will take some time to figure out. **Jones** suggested that engaging people remotely does provide an option as well. **Hanson** suggested that, as the additional simulations are done, perhaps we could meet via webinar to follow along. **Leaf** indicated that the additional simulation work will take approximately six weeks or less.

Mambretti suggested a thorough re-review of data we have used in past in the index development. Some of the TPWD data was censored in the past because of questionable IDs. Can we revisit the value of those data streams to contribute to the index before we move forward with any additional simulations? **VanderKooy** agreed that all the data will need to be vetted for next assessment. We cannot change the inputs at this time given the current schedule. Per **Schueller**, **Mambretti's** idea would result in a whole new benchmark assessment and would be done prior to next assessment. It needs to be thought of as a turn of the crank where we do not change the inputs. **Hanson** wondered if there was a point where the turn of the crank allows evaluation of the other datasets and how they may or may not correlate to the original indices. **Jones** heard that that would occur as a benchmark in a few years. **Schueller** does not update the other datasets in between assessments if it is not used IN the benchmark. An annual update is not likely with the current timeline between the two Commissions assessment needs. **Butterworth**: "In order to develop another index, you need to run it through a full benchmark again but you could look at the results from another survey just to see if it trends in the same direction."

VanderKooy will assemble the report for this workshop and distribute to the whole group. He will facilitate the process on getting the mechanisms started by coordinating with the states and industry in anticipation of an operational MP proposal for the MAC meeting in October. The Commission meeting will be in Mississippi and **VanderKooy** will let everyone know when details are final. Thanks to all those who did come and especially to those who stayed. Kudos to **Hanson**.

VanderKooy explained the Commission's travel procedures and asked everyone who intended to submit, to do so quickly.

Public Comment
There were no comments received as the public was at the table and there were no participants on the webinar portion.

Adjourn
With no further business or discussion, the workshop closed at 11:35 a.m. Friday.

Appendix A - July paper_fin

Extensions of the Application of MSE to Gulf Menhaden

Doug S. Butterworth and Rebecca A. Rademeyer

July 2019

Summary

This document provides extends the MSE simulation-testing process suggested for Gulf Menhaden in Rademeyer and Butterworth (2019) to make specific suggestions for a Management Procedure (MP) which would provide a basis to impose catch limitations on the fishery, though only when survey indices suggest that resource abundance has dropped further than is desirable. In particular, a reasonably large number of robustness tests has been developed to check if the MP (harvest control rule, which would set a catch limit based on recent values of abundance survey indices) provides robust performance (particularly as regards safeguarding the resource from undue depletion) in the face of assessment and other uncertainties. It transpires that for the current range of annual catch sizes, such a rule is really necessary only in the future circumstances of increasing natural mortality or a period of poor recruitment. A tuning criterion is put forward as a basis to compare the performances of different MPs (essentially different choices for the parameters of the harvest control rule). Finally, a suggestion is made for the value of one of these parameters (that of a composite survey abundance index) below which a catch limit would be imposed; higher values for this choice would lead to a greater frequency of (unnecessarily) imposing a catch limit, whereas lower values result in smaller values for lowest level of egg production expected and higher values of the average annual variability in landings. The composite survey index has been above the threshold value suggested for the last ten years.

Introduction

This document provides extensions to the simulation-testing process of Management Strategy Evaluation (MSE) suggested in Rademeyer and Butterworth (2019) for Gulf Menhaden. The objective is to develop a "Management Procedure" (MP) that provides a basis to impose catch limitations on the fishery, though only when survey indices suggest that abundance has dropped further than is desirable. The reasons for the desirability of such an approach are set out in Rademeyer and Butterworth (2019). The document first develops a basis to test the robustness of such an MP to alternative possible dynamics for the Gulf Menhaden resource and its associated fishery. It then proceeds to provide the results of such tests for alternative possible MPs (essentially different harvest control rules), and provides suggestions on an appropriate trade-off choice on which to base the final selection of an MP.

Methods

The Base Case Operating Model (OM) taken forward here to reflect the dynamics of the Gulf Menhaden population as a basis for MP testing mimics the BAM Base Model developed for the assessment of this resource (SEDAR, 2018).

<u>The projections</u>

Key aspects of the 20-year projections conducted are as follows, with full details (including some exceptions to the broad statements made below) set out in Appendix A. Note that the second, third and last of the bullets below reflect adjustments from procedures followed in Rademeyer and Butterworth (2019).

- Unless otherwise specified, future dynamics are the same as for the BAM Base Model assessment.
- Future annual landings are drawn at random, with replacement, from the 2000-2017 values. The landings in 2018 are taken to be 525 635 mt. Since already half of 2019 has passed, a catch drawn at random from 2000-2017 is assumed for 2019 and then projected forward on this same basis, except when overridden by the harvest control rule, from 2020 to 2039.
- A maximum full fishing mortality (Fmax, taken as 105% of the estimated historical maximum full fishing mortality) is imposed to avoid unrealistic values, i.e. instances where the low size of the resource makes it unlikely the future intended catch could be taken, so that this is overridden by a value corresponding to Fmax. If future fishing mortality is computed to be above Fmax, then the selectivity for that year for age 1 is changed to 0.8 and the fishing mortality is recomputed. If this recomputed fishing mortality is still above Fmax, the landing is recalculated assuming an apical fishing mortality of Fmax (and the "widened" selectivity). The choice of 0.8 (increased from the 0.6 suggested in Rademeyer and Butterworth (2019)) has been made so as to reduce the chance that the resource is "protected" from undue depletion through inability to make the intended catch, rather than by the management rule (MP), and hence provides a more stringent test of the efficacy of that rule.
- A hockey-stick is assumed for the stock-recruitment curve, with the break taken as SSB=1.8×10^6 (in billions of eggs)[1] – see Figure 1.
- Future recruitment residuals are drawn at random, with replacement, from the 1978-2017 model estimated residuals.
- Future survey results are computed assuming log-normal observation error, with standard deviation computed from past (2005+) model estimated error. The selectivity and catchability values are taken to be as estimated for the BAM Base Model. Auto-correlation has been included in the future for the seine index, with the autocorrelation coefficient as estimated in the conditioning[2] of the OM concerned.

Robustness tests

Robustness tests have been developed over recent months in collaboration with a technical group consisting of David Chagaris, Peter Himchak, Robert Leaf, Genny Nesslage and Amy Schueller. These tests are listed in Table 1, and fall into two categories.

a) OMs considered to reflect alternative plausible realities to the Base Case OM, for which any MP considered for implementation must evidence reasonably robust performance (Type A).
b) Other OMs whose plausibility is low at best, but which have been included more with a view to check how far the MPs considered can be "pushed" before they provide inadequate performance) (Type B).

The Management Procedure considered

[1] For convenience all future reference to numbers of eggs ("SSB") will be in units of million billions, so that the break value indicated here becomes 1.8.

[2] There is no indication of auto-correlation in recruitment. For the gillnet index, auto-correlation varies substantially depending on the relative weighting assigned to the two indices, and could be appreciably negative, suggesting (questionably) enhanced precision of the index, so that it was decided to set it to zero when projecting.

The MP considered is empirical. It overrides and reduces a landing drawn from the historical set only if the value of a combined abundance index (J_y for year y) falls below a threshold level ($J_{threshold}$) specified for that index. If the threshold is breached, a TAC is set based of the value of this combined index, which is a weighted average of the gill net and seine indices, i.e.:

If $J_y < J_{threshold}$:

$$TAC_{y+1} = \gamma J_y \quad (1)$$

Figure 2 illustrates the rule for a initial choice of control parameter parameter values ($J_{threshold} = 0.8$ and $\gamma = 500$) for this "harvest control rule", and also plots historical values of the combined index. More details are given in Appendix B.

The performance of the MP is reported in terms of a number of performance statistics, which are listed and defined in Appendix C.

Results

Results for conditioning (i.e. fitting the BAM for) the Base Case and Robustness test OMs are shown in Appendix D. Note that results are required only for those tests which involve historical (and not projection only) changes to the Base Case, so that the OM has to be refitted. Throughout the conditioning appears satisfactory, with no indications of systematic lack of fit to the abundance indices.

Figure 3 provides a summary of certain key performance statistics for the Base Case and all the Robustness test OMs, indicating the differences in performance with and without the baseline MP (harvest control rule).

It is evident from Figure 3 that for the more plausible (Type A) OMs, only in the cases of Robustness tests 1.5 (increasing natural mortality M in the future) and 4.1 (a period of decreased recruitment in the future) is there any need for some restriction along the lines of a harvest control rule to counter undesirable depletion of the resource through harvesting. Hence it is only for these OMs that projection plots showing the differences in performance with and without the baseline MP in place are shown in the main text (Figure 4 and Table 2). The corresponding plots, together with a Table of performance statistics, for the rest of the OMs are provided in Appendix E.

MP variants

Appendix F shows results for the Base Case OM, and the 1.5 and 4.1 Robustness tests, for changed values of the three control parameters of the Baseline MP ($J_{threshold}$, γ and p). It is evident that as the value of p is increased, there is a trade-off between an increase in the lowest landing, but a decrease in the lowest egg production (denoted by SSB) value to be expected (this occurs because with a larger value of p, there is a greater delay in an adequate response to recent poor values for the resource indices). For further evaluations, the value of p was set to 2 to reflect a reasonable choice for this trade-off.

For readier comparison of results, the choice of the other two control parameter values ($J_{threshold}$ and γ) was made by fixing the value of $J_{threshold}$ and then tuning the corresponding value of γ so that the median lowest egg production (denoted by SSB) for the 4.1 Robustness test was equal to 1.0. One cannot expect to achieve the same minimum abundance in the Robustness tests as for the Base Case OM, as they reflect less net resource productivity; note that Figure 5 indicates that in the absence of landings, the lowest resource SSB on projection is a little over 3, whereas for both the 1.5 and 4.1 Robustness tests this becomes only a little larger than 2. The choice of the value of median SSB = 1.0 as the criterion for tuning was that it similarly achieves a median lowest SSB on projection under harvest that is about 1.0 less than that under harvest for the Base Case OM (see Figure 5; in addition, projections are shown in Figure 6).

For reasons discussed below the tuned MP with $J_{threshold}$ = 0.9 seems to provide the best trade-offs, and is therefore "advocated". Further results for projections under this MP for the Base Case OM, and for the 1.5 and 4.1 Robustness tests, are shown in Figures 7 and 8, with performance statistics reported in Table 3.

As a sensitivity, the replacement of a linear by a quadratic harvest control rule was explored, i.e.:

If $J_y < J_{threshold}$:

$$TAC_{y+1} = \gamma {J_y}^2 \qquad (2)$$

with $J_{threshold}$ fixed at 0.9 and γ tuned to 691 so that the median lowest egg production (denoted by SSB) for the 4.1 Robustness test was equal to 1.0.

This improves the lower 5%-ile for the lowest SSB for the two key Robustness tests, though not for the Base Case OM; but this is at the expense of lower lowest landings and higher average annual landings variability (AAV) (see Figure 9 and Table 3). As the benefits of this change appear outweighed by the disadvantages, the choice of a linear rule would seem to be preferred.

Discussion

Figure 5 provides information on the trade-offs involved in making an appropriate choice for the value of the $J_{threshold}$ control parameter. Once $J_{threshold}$ exceeds 0.9, the values of performance statistics shown stabilise, so there seems no advantage in setting this value higher in circumstances where that would have the adverse consequence of the catch limit needing to be imposed more frequently. On the other hand, when $J_{threshold}$ is set lower than 0.9, the lowest landing anticipated drops and AAV increases. This suggests that the choice of an MP with $J_{threshold}$ = 0.9, $\gamma = 500$ thousand mt and p = 2 could be appropriate. Note that the composite index has not fallen below this 0.9 value in the last ten years (see Figure 2).

These values do, however, follow given the tuning choice of a median lowest SSB of 1.0 for the 4.1 Robustness test. While a basis for choosing this value is offered above, a more or less conservative MP could be obtained by increasing or decreasing this choice for the tuning value.

References

Rademeyer, R.A. and Butterworth, D.S. 2019. An initial illustrative example of the application of MSE to Gulf Menhaden to address Issues related to MSC certification and Ecosystem-related Reference Points. Document circulated for internal discussions, January 2019. (17pp)

SEDAR. 2018. SEDAR 63 – Gulf Menhaden Stock Assessment Report. SEDAR, North Charleston SC. 352 pp. available online at: http://sedarweb.org/sedar-63

Table 1: List of the robustness tests used in MP testing. Note that "No refitting" means that the test involves changes in the future only. Type A OMs are considered to reflect alternative plausible realities to the Base Case OM, while the plausibility of Type B OMs is low at best, but these OMs have been included more with a view to check how far the MPs considered can be "pushed" before they provide inadequate performance.

	Base Case	Robustness	No refitting	Type
1. Alternative choices for *M*				
1.1		M'(a)=1.2		A
1.2		M'(a)=M(a)*exp(-0.1(a-2))		A
1.3	Lorenzen mortality vector	M(4+)=1.67		A
1.4		M increases linearly by 40% over next 20 years	x	B
1.5		M increases linearly by 20% over next 20 years	x	A
2. Alternative catch selectivity function				
2.1	S(3) = S(4+) = 0.87	S(3) = S(4+) = 1.0		A
2.2		S(3) = S(4+) = 0.74		A
2.3	S(1) in future as estimated in past	S(1) in future, double that estimated in the past	x	B
3. Indices				
3.1	Linear relationship to abundance: I = q*B	sqrt relationship to abundance I =q*sqrt(B)		A
3.2	Weighting: 4:1 gillnet to seine	Weighting: 1:1 gillnet to seine		A
3.3		Observation error = 0.2	x	A
3.4	Observation error = 0.11	Observation error = 0.3	x	B
3.5		Observation error = 0.5	x	B
3.6	Flat 2+ gillnet selectivity in the future	Increasing 2+ sel. slope over the next 20 years (to 0.4 age 4 in 20yrs)	x	B
4. Period of future poor recruitment				
4.1	Future rec. drawn at random from past values	Five (2020-2024) years of bad recruitments (50%)	x	A
5. Alternative stock-recruitment function				
5.1	Hockey-stick, hinge-point=1.8 billion eggs	Hockey-stick, hinge-point=2.2 billion eggs	x	A
6. Under-reporting of future catches (which is not noticed)				
6.1	Future catches=TAC	Future catches = 1.1TAC (presence of these IUU catches is not realised)	x	A
7. Maximal possible fishing mortality				
7.1	Fmax for projections = 1.05*Fmax historical	Fmax for projections = 1.20*Fmax historical	x	A

Table 2: Performance statistics for the Base Case OM and Robustness tests 1.5 and 4.1 with and without the management rule (Baseline MP).

Performance statistics	Base Case			Robustness 1.5			Robustness 4.1		
NO RULE	Median	10	90	Median	10	90	Median	10	90
Related to catch									
Average landing 2020-2039	494.7	479.2	517.5	484.2	387.8	509.3	371.9	169.9	500.7
Av landing no rule	494.7	479.2	517.5	484.2	387.8	509.3	371.9	169.9	500.7
Av landing with rule	-			-			-		
Lowest landing (2020-2039)	379.9	379.9	425.6	379.9	72.2	400.7	181.9	41.0	379.9
2020 landing	473.7	400.7	590.8	473.7	400.7	590.8	473.7	400.7	590.8
Related to abundance									
Egg(2020)	3.32	2.73	4.10	3.32	2.73	4.10	3.32	2.73	4.10
Egg(2040)	3.17	2.21	4.04	0.78	0.08	1.97	0.59	0.10	3.80
Egg lowest (2020-2039)	2.00	1.44	2.38	0.76	0.08	1.50	0.30	0.05	1.06
Prob Egg(2040) lowest	6			63			9		
Related to catch variability									
AAV 2020-2039	0.15	0.12	0.19	0.16	0.12	0.21	0.20	0.14	0.27
AAV with rule	-			-			-		
Other									
Fraction years rule applied	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
True negative	0.0			0.0			0.0		
Fals negative	0.0			0.0			0.0		
False positive	0.0			0.0			0.0		
True positive	100.0			100.0			100.0		
Prob rule in 2020	0			0			0		
Fraction years Hit Fmax	0	0	0.15	0.35	0.1	0.65	0.85	0.25	0.9
Hit Fmax, landings not taken	0	0	0	0.05	0	0.35	0.65	0	0.85
WITH BASELINE MP	Median	10	90	Median	10	90	Median	10	90
Related to catch									
Average landing 2020-2039	483.8	448.6	511.0	423.9	368.9	466.1	382.4	204.5	459.3
Av landing no rule	494.8	477.8	519.1	497.8	477.0	519.6	498.1	467.3	529.6
Av landing with rule	344.4	0.0	389.5	321.0	234.5	360.0	277.2	145.9	389.0
Lowest landing (2020-2039)	364.2	272.8	400.7	246.9	138.4	336.1	176.5	43.3	300.1
2020 landing	473.7	400.7	590.8	473.7	400.7	590.8	473.7	400.7	590.8
Related to abundance									
Egg(2020)	3.32	2.73	4.10	3.32	2.73	4.10	3.32	2.73	4.10
Egg(2040)	3.19	2.37	4.04	1.59	0.87	2.33	3.16	2.18	4.01
Egg lowest (2020-2039)	2.01	1.48	2.38	1.23	0.67	1.66	0.51	0.14	1.31
Prob Egg(2040) lowest	5			25			0		
Related to catch variability									
AAV 2020-2039	0.16	0.13	0.20	0.17	0.12	0.23	0.18	0.14	0.24
AAV with rule	0.15	0.00	0.26	0.17	0.10	0.24	0.16	0.10	0.27
Other									
Fraction years rule applied	0.10	0.00	0.25	0.40	0.20	0.60	0.53	0.35	0.80
True negative	10.1			43.7			55.5		
Fals negative	0.8			1.4			0.4		
False positive	6.8			12.5			5.8		
True positive	82.4			42.5			38.3		
Prob rule in 2020	0			0			0		
Fraction years Hit Fmax	0	0	0.1	0.1	0	0.2	0.3	0.1	0.45
Hit Fmax, landings not taken	0	0	0	0	0	0	0.05	0	0.15

Table 3: Performance statistics for the Base Case OM and Robustness tests 1.5 and 4.1 without the management rule ("No rule"), with the advocated MP ("MP_2_0.9") and with the quadratic MP ("MP_2_0.9_quad").

	Base Case									Robustness test 1.5									Robustness test 4.1								
	No rule			MP_2_0.9			MP_2_0.9_quad			No rule			MP_2_0.9			MP_2_0.9_quad			No rule			MP_2_0.9			MP_2_0.9_quad		
	Median	10	90	Median	10	90	Median	10	90	Median	10	90	Median	10	90	Median	10	90	Median	10	90	Median	10	90	Median	10	90
Related to catch																											
Average landing 2020-2039	494.7	479.2	517.5	472.3	440.8	502.4	479.8	445.4	509.6	363.3	288.9	432.8	414.0	364.5	455.0	417.4	356.3	452.6	371.9	169.9	500.7	395.9	323.8	439.5	393.9	344.6	438.0
Av landing no rule	494.7	479.2	517.5	496.3	476.4	517.4	495.0	476.4	519.4	363.3	288.9	432.8	501.1	465.7	530.0	500.5	462.5	527.8	371.9	169.9	500.7	495.4	470.3	526.7	494.4	470.0	521.3
Av landing with rule	-			390.0	330.1	426.7	415.6	325.4	496.6	-			347.0	271.5	388.5	344.8	255.1	406.6	-			290.9	221.3	352.8	277.2	202.6	363.5
Lowest landing (2020-2039)	379.9	379.9	425.6	347.5	261.6	397.2	322.2	190.0	400.7	14.6	2.8	69.3	236.1	140.2	318.3	167.9	73.2	265.3	181.9	41.0	379.9	174.7	84.0	254.2	97.3	31.9	191.3
2020 landing	473.7	400.7	590.8	473.7	400.7	590.8	473.7	400.7	590.8	473.7	400.7	590.8	473.7	400.7	590.8	473.7	400.7	590.8	473.7	400.7	590.8	473.7	400.7	590.8	473.7	400.7	590.8
Related to abundance																											
Egg(2020)	3.32	2.73	4.10	3.32	2.73	4.10	3.32	2.73	4.10	3.32	2.73	4.10	3.32	2.73	4.10	3.32	2.73	4.10	3.32	2.73	4.10	3.32	2.73	4.10	3.32	2.73	4.10
Egg(2040)	3.17	2.21	4.04	3.20	2.37	4.07	3.20	2.37	4.05	0.01	0.00	0.07	1.63	0.89	2.39	1.76	1.01	2.40	0.59	0.10	3.80	3.19	2.37	4.07	3.20	2.37	4.05
Egg lowest (2020-2039)	2.00	1.44	2.38	2.02	1.60	2.40	2.01	1.52	2.39	0.01	0.00	0.07	1.32	0.77	1.71	1.31	0.88	1.67	0.30	0.05	1.06	1.00	0.48	1.50	1.00	0.60	1.51
Prob Egg(2040) lowest	6			5			5			100	100	100	24			19			9	9	9	0			0		
Related to catch variability																											
AAV 2020-2039	0.15	0.12	0.19	0.17	0.12	0.21	0.18	0.12	0.25	0.22	0.17	0.28	0.17	0.12	0.25	0.25	0.17	0.35	0.20	0.14	0.27	0.21	0.16	0.27	0.30	0.21	0.42
AAV with rule	-			0.16	0.05	0.24	0.20	0.05	0.43	-			0.17	0.12	0.24	0.29	0.19	0.46	-			0.23	0.15	0.35	0.41	0.25	0.68
Other																			0								
Fraction years rule applied	0.00	0.00	0.00	0.20	0.05	0.40	0.20	0.05	0.40	0.00	0.00	0.00	0.60	0.35	0.70	0.60	0.35	0.70	0.00	0.00	0.00	0.50	0.35	0.65	0.50	0.30	0.65
True negative	0.0			19.4			20.8			0.0			58.0			58.4			0.0			48.7			47.1		
Fals negative	0.0			2.6			2.3			0.0			1.5			1.4			0.0			1.9			2.1		
False positive	0.0			11.4			11.6			0.0			11.6			11.9			0.0			10.0			10.3		
True positive	100.0			66.8			65.4			100.0			29.0			28.4			100.0			39.5			40.6		
Prob rule in 2020	0			0			0			0			0	0	0	0	0	0	0			0	0	0	0	0	0
Fraction years Hit Fmax	0	0	0.15	0	0	0.05	0	0	0.05	0.6	0.4	0.75	0.05	0	0.1	0.05	0	0.15	0.85	0.25	0.9	0.05	0	0.15	0.05	0	0.1
Hit Fmax, landings not taken	0	0	0	0	0	0	0	0	0	0.4	0.25	0.55	0	0	0	0	0	0	0.65	0	0.85	0	0	0	0	0	0

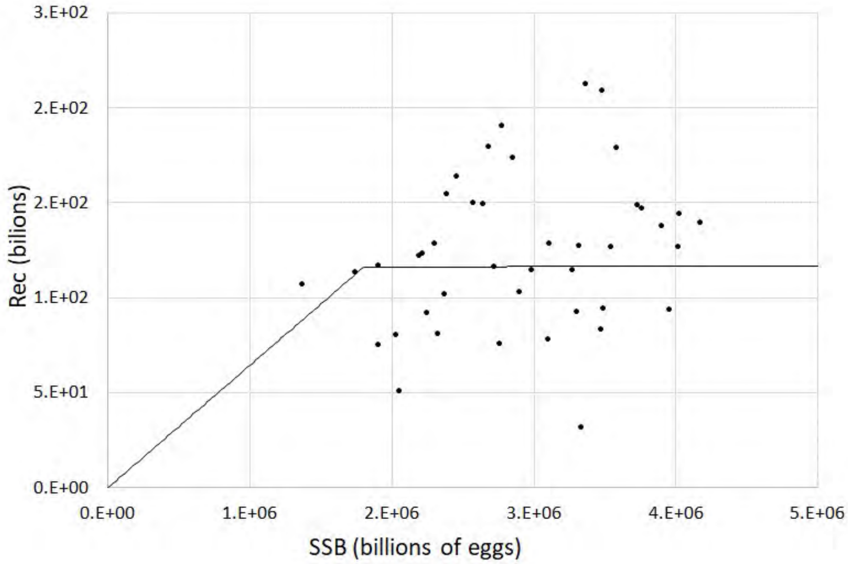


Figure 1: Hockey-stick stock recruitment curve for Gulf Menhaden which is used to compute projected recruitment. The data points are those estimated in the BAM Base Model.

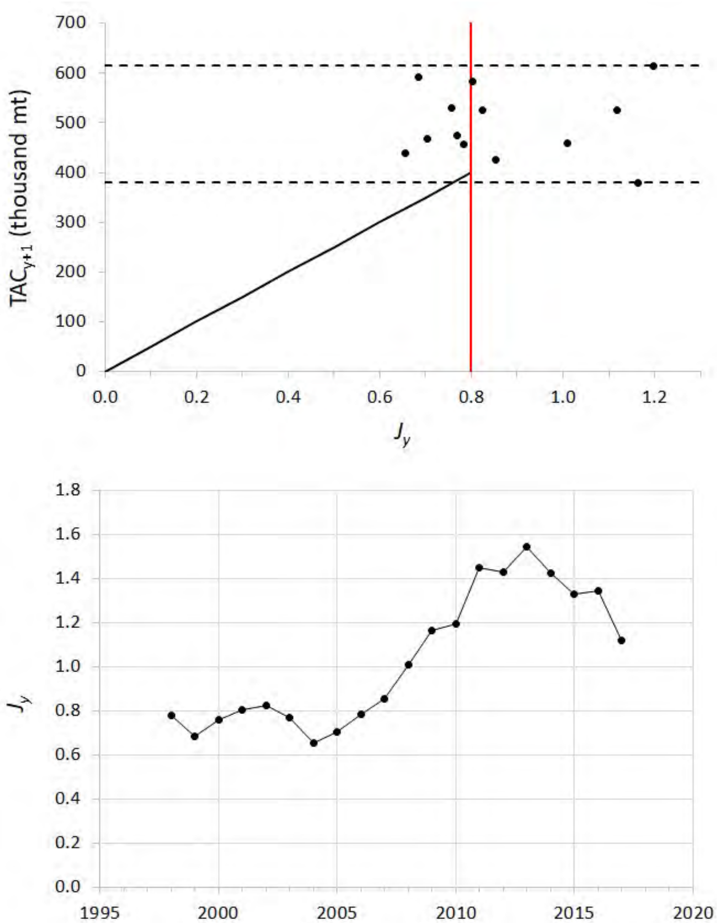


Figure 2: Top plot: Illustration of the management rule (MP) for set control parameter values considered in the example for which results are reported. The horizontal dash lines show the 2000-2017 minimum

and maximum landing values. The historical (1999-2017) J_y vs TAC_{y+1} are shown as black dots. Bottom plot: Historical combined index J_y values.

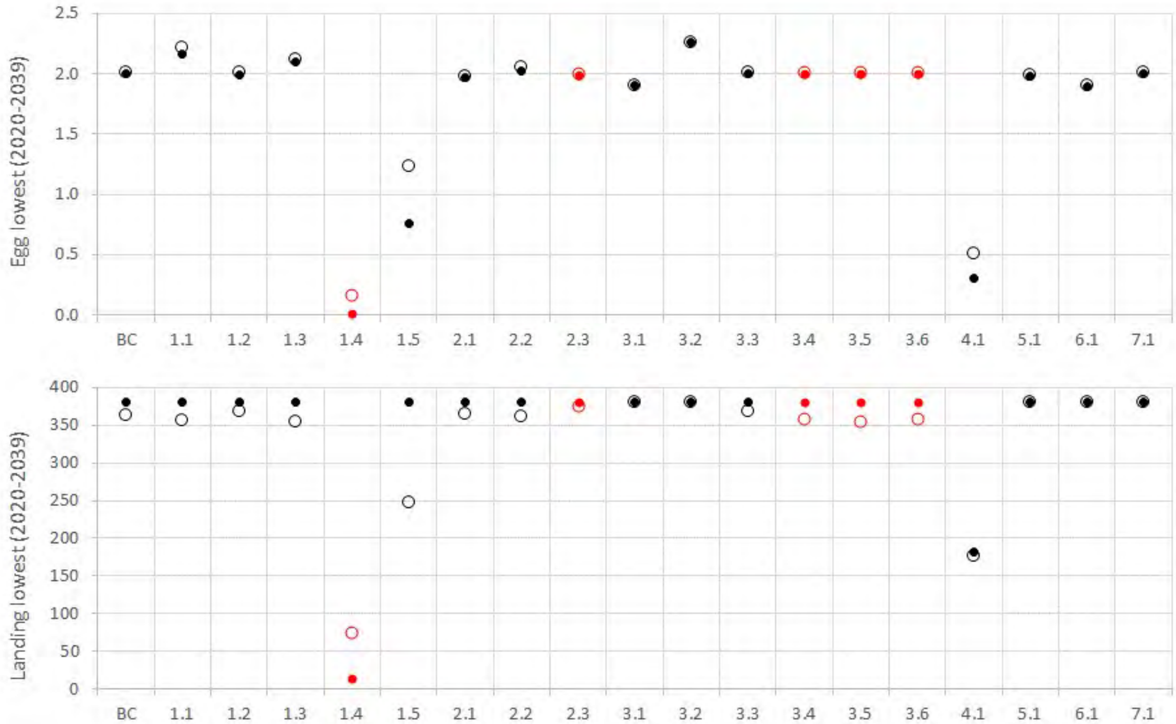


Figure 3: Median lowest egg production and landing values over the 2020-2039 projection period for each of the Base Case and Robustness test OMs without (full circles) and with (open circles) the Baseline MP. Type B OMs are shown in red.

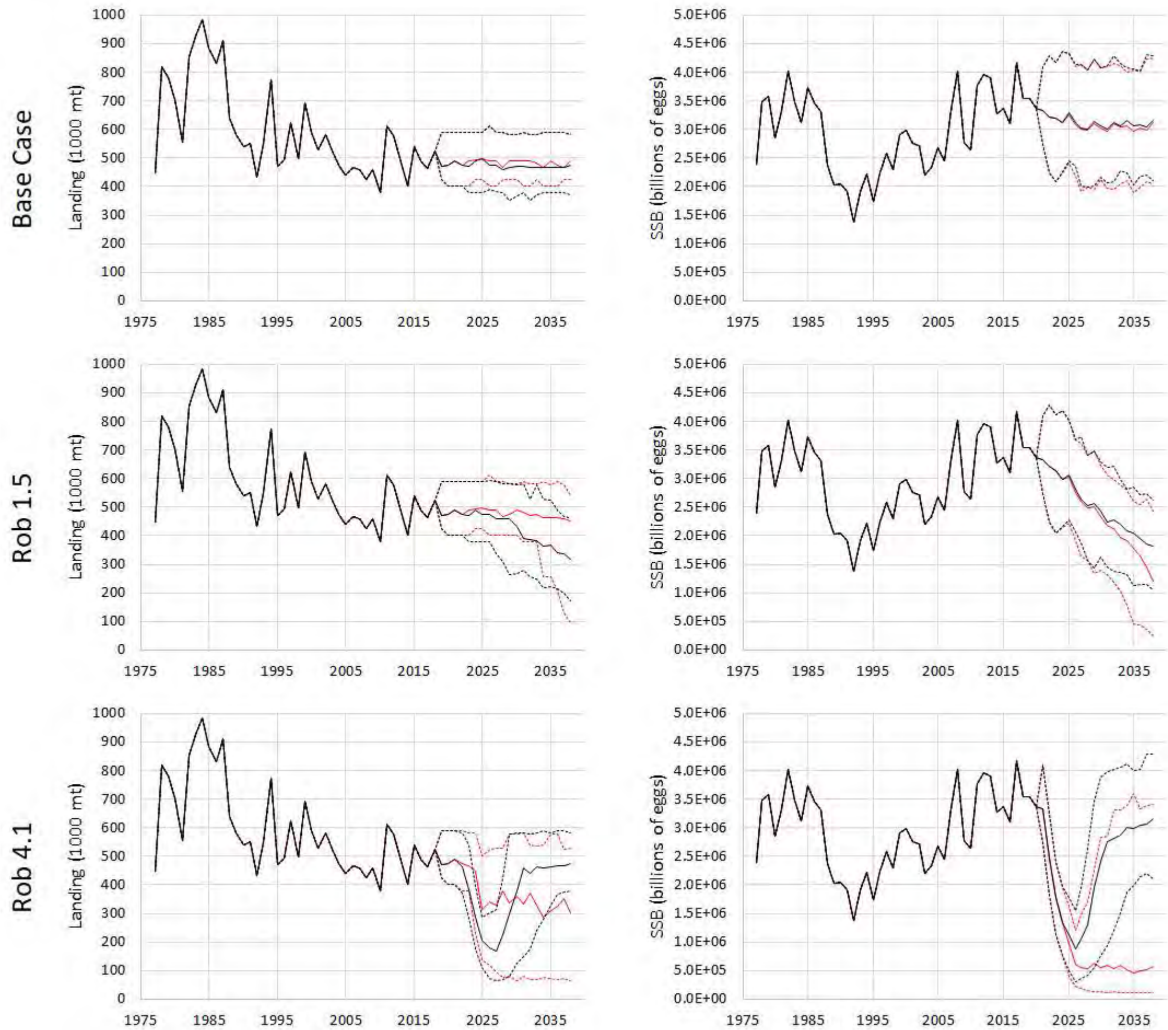


Figure 4: Historical estimates and projected 20-year median with 10%- and 90%-iles for a series of quantities for the **Base Case** OM and **Robustness tests 1.5 and 4.1**, without (red lines) and with the management rule (Baseline MP, black lines)

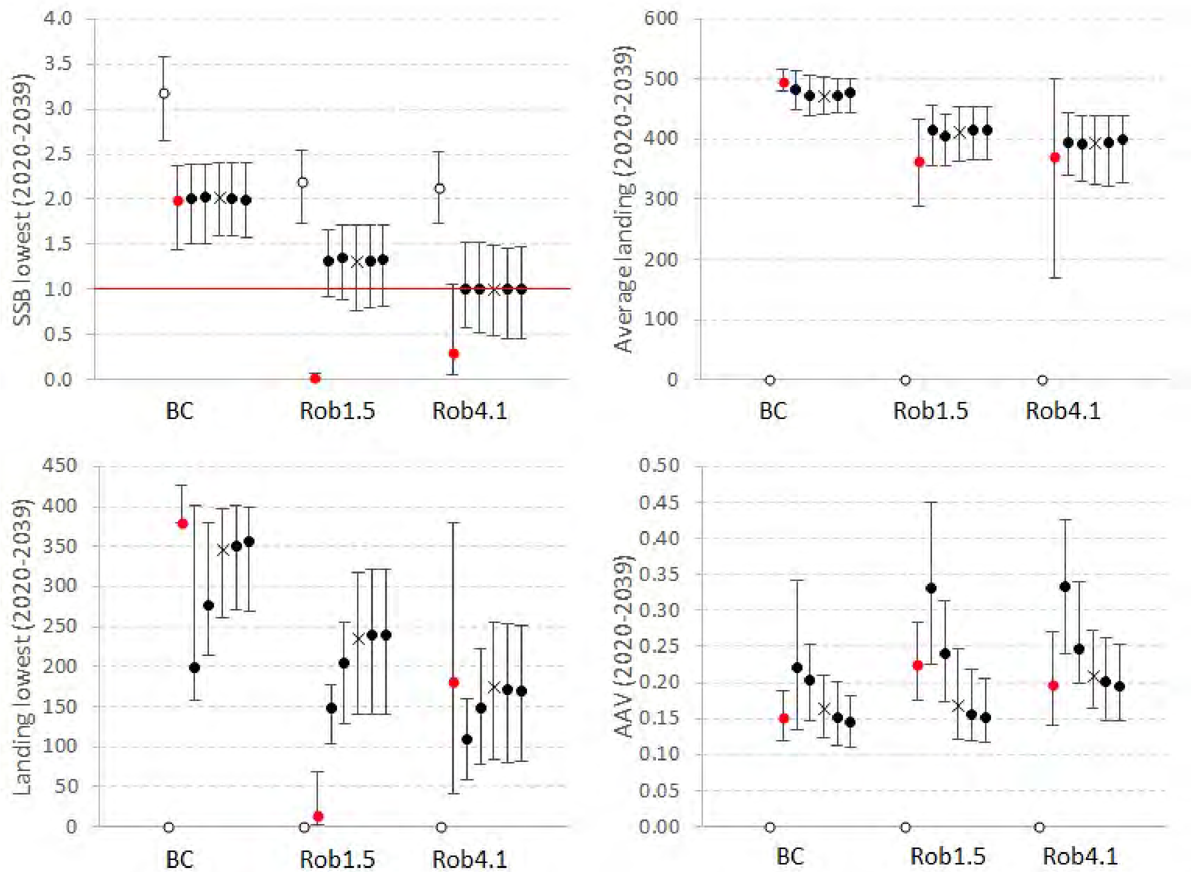


Figure 5: Performance statistics for no landings (for SSB lowest only) (open dot), no harvest control rule (red dot) and MP variants with p = 2 for varying the value of the $J_{threshold}$ control parameter. The γ control parameter value is tuned so that the median SSB for Robustness test 4.1 (poor future recruitment trial) is equal to 1 (shown by the horizontal red line). Results (median with 10%- and 90%-iles) are shown for the Base Case OM and the Robustness tests 1.5 and 4.1. The ($J_{threshold}$;) combinations shown are (0.7 ; 293); (0.8 ; 400); (0.9 ; 500) – advocated and shown by crosses; (1.0 ; 505) and (1.1 ; 500).

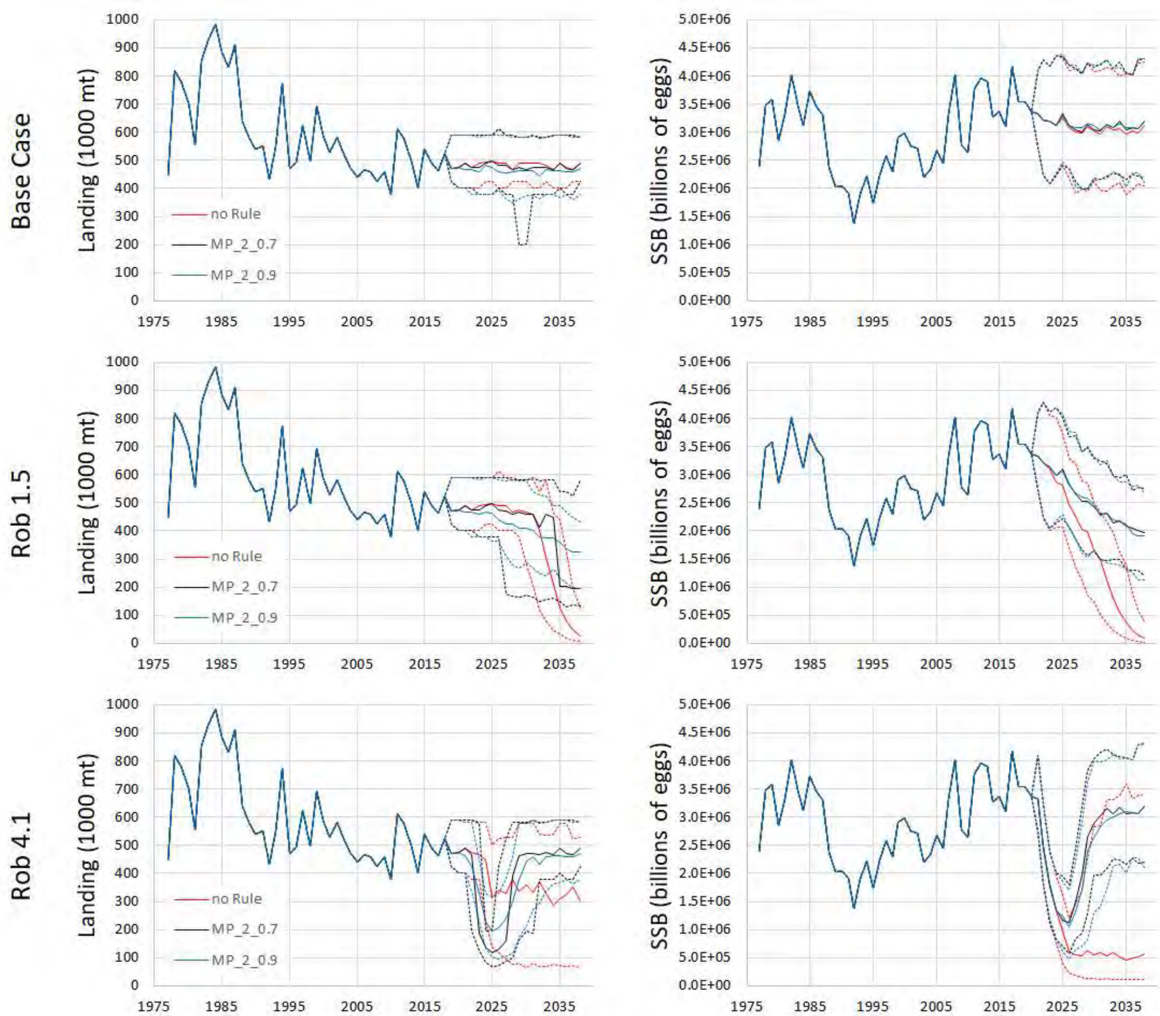


Figure 6: Median (full line) with 10%- and 90%-iles for projected landings and SSB for no harvest control rule and MP variants with p = 2 for $J_{threshold}$ = 0.7 and = 0.9, for the Base Case OM and Robustness tests 1.5 and 4.1.

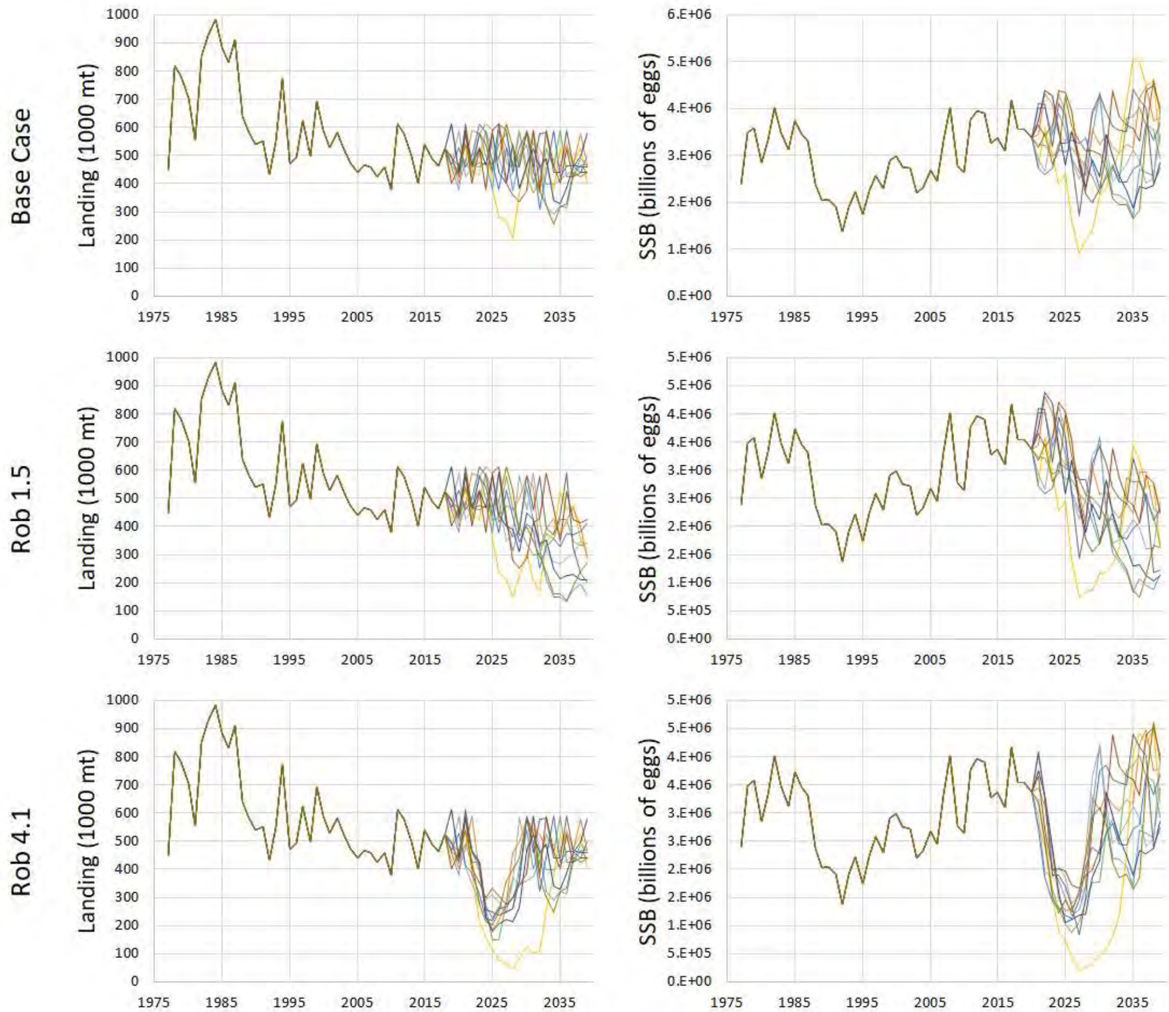


Figure 7: Worm plots for projected landings and SSB for the advocated MP variant with p = 2 for $J_{threshold} = 0.9$, for the Base Case and Robustness tests 1.5 and 4.1.

Figure 8a: Historical estimates and projected 20-year median with 10%- and 90%-iles for a series of quantities for the Base Case OM, without (red lines) and with (black lines) the advocated management rule.

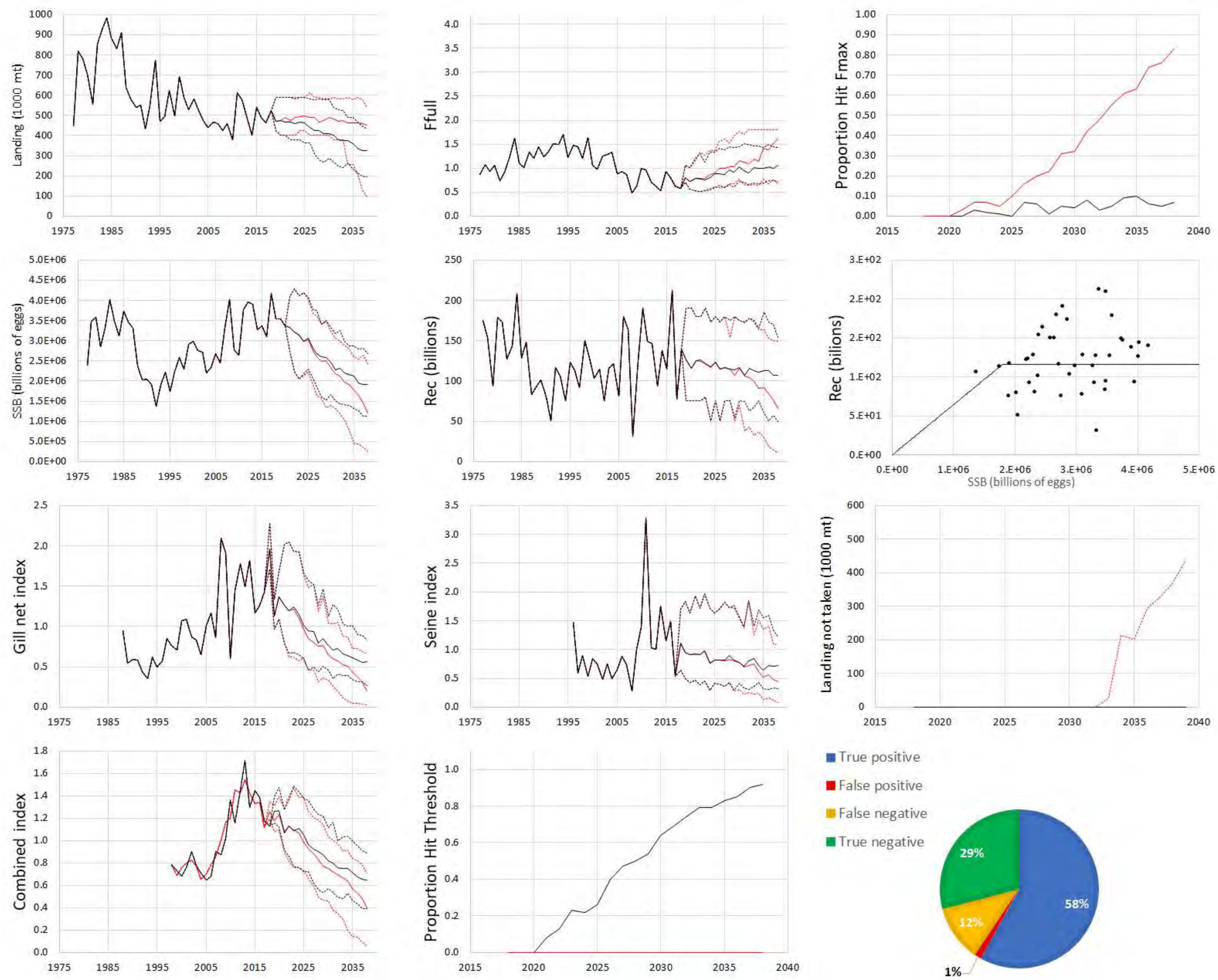


Figure 8b: Historical estimates and projected 20-year median with 10%- and 90%-iles for a series of quantities for Robustness test 1.5, without (red lines) and with (black lines) the advocated management rule.

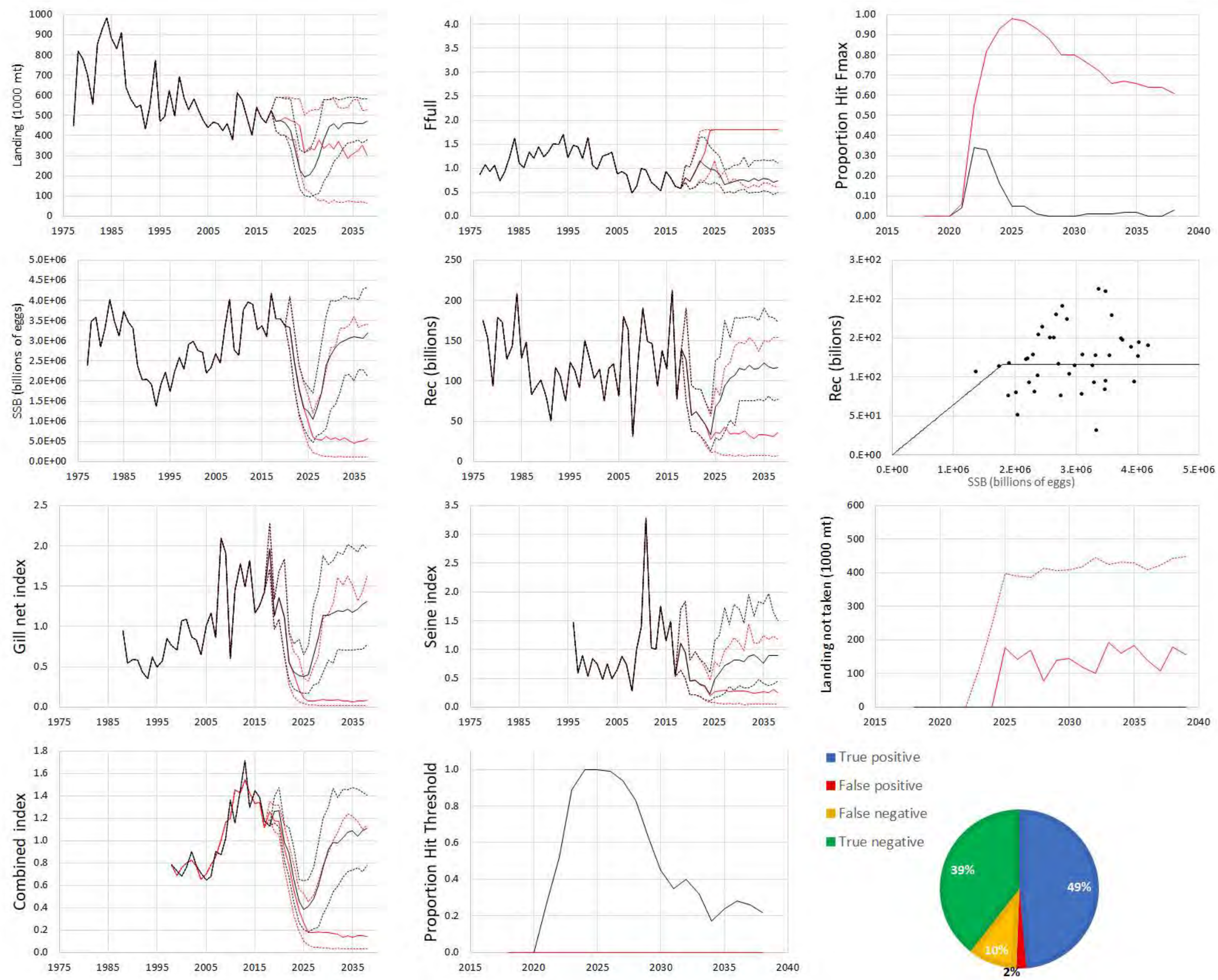


Figure 8c: Historical estimates and projected 20-year median with 10%- and 90%-iles for a series of quantities for Robustness test 4.1, without (red lines) and with (black lines) the advocated management rule.

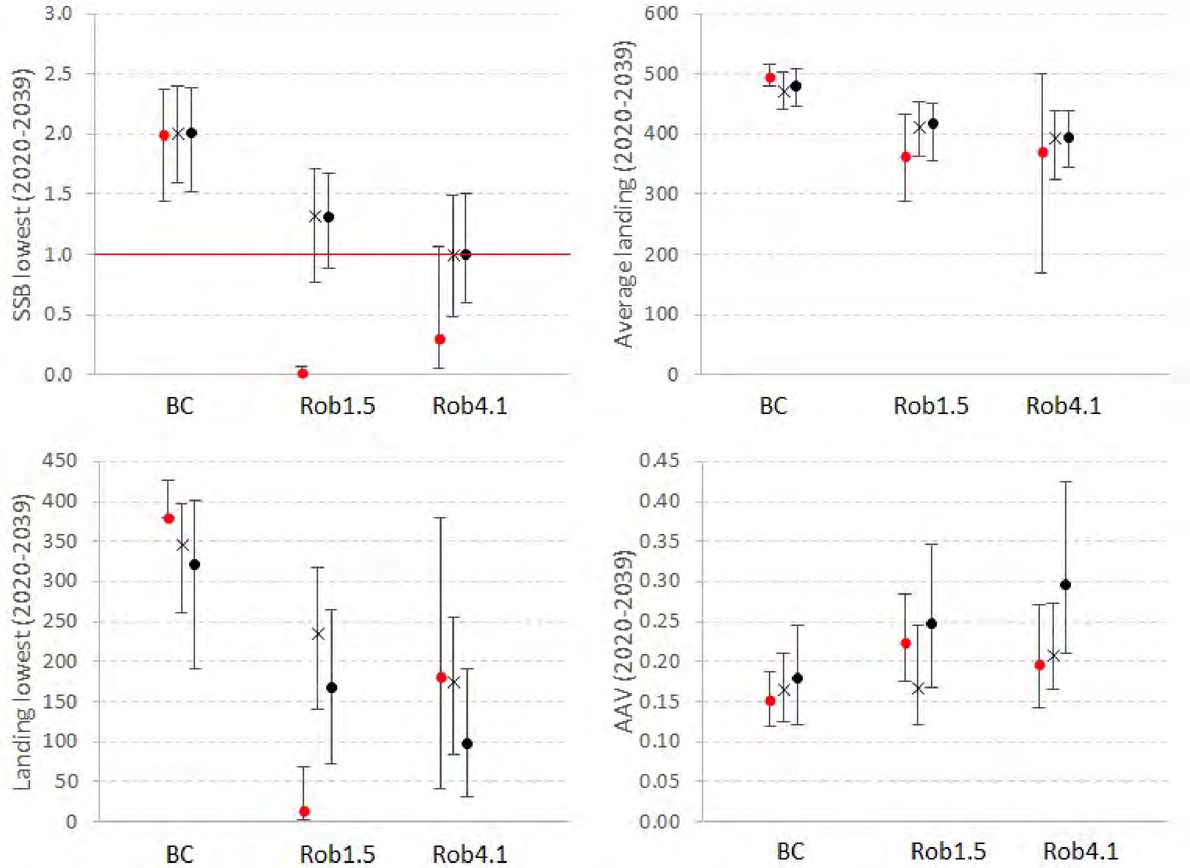


Figure 9: Performance statistics for a) no harvest control rule (red dot), b) the advocated MP variant with $p = 2$ and $J_{threshol}$ = 0.9 (cross) and c) and MP variant with a quadratic instead of linear function (i.e. $TAC_{y+1} = \gamma {J_y}^2$), with $p = 2$ and $J_{threshold}$ = 0.9. As for the linear MP, the γ control parameter value is tuned so that the median SSB for the 4.1 Robustness test (poor future recruitment trial) is equal to 1 (shown by the horizontal red line). Results (median with 10%- and 90%-iles) are shown for the Base Case OM and the 1.5 and 4.1 Robustness tests.

Appendix A – Projection methodology details

Projections into the future under a specific management rule (MP) are performed using the following steps.

Step 1: Begin-year numbers at age

The components of the numbers-at-age vector at the start of 2018 ($N_{2018,a}$: *a* =1,…, *m* – where *m* is a plus-group) are obtained from the MLEs for an assessment of the resource. The assessment used here is the BAM Base model.

Step 2: Annual landings

For 2018, $L_{2018} = 525\ 635$ mt. (A.1)

From 2019 onwards:

L_y is drawn at random, with replacement, from the observed 2000-2017 landings.

From 2020, if the combined abundance index (see equation B2 In Appendix B) for year *y*-1 is below the threshold value, then a TAC applies to year *y* is computed using the MP (harvest control rule) (see equation (1) of the main text and Appendix B).

Step 3: Landings-at-age (by number)

The $L_{y,a}$ values are obtained under the assumption that the commercial selectivity function (S_a) estimated for the most recent period in the BAM Base Model (1996+) continues in the future. The full fishing mortality F_y is solved iteratively to achieve the annual landing by mass:

$$L_y = \sum_{a=1}^{m} w_a^{mid} N_{y,a} S_a F_y \left(1 - e^{-Z_{y,a}}\right) / Z_{y,a} \quad \text{(A.2)}$$

where

w_a^{mid} is the time invariant weight-at-age in the middle of the year,

$N_{y,a}$ is the number-at-age vector for age *a* at the start of year *y* (with *m* the plus group),

and

$Z_{y,a} = F_y S_a + M_a$ is the total mortality-at-age vector for age *a* and year *y*.

M_a is the natural mortality-at-age *a* (input).

The numbers-at-age can then be computed for the beginning of the following year (*y*+1):

$$N_{y+1,1} = R_{y+1} \quad \text{(A.3)}$$

$$N_{y+1,a+1} = N_{y,a} e^{-Z_{y,a}} \quad \text{for } 1 \le a \le m-2 \quad \text{(A.4)}$$

$$N_{y+1,m} = N_{y,m-1} e^{-Z_{y,m-1}} + N_{y,m} e^{-Z_{y,m}} \quad \text{(A.5)}$$

If the intended landing is such that the apical fishing mortality (that at the age at which selectivity is 1) exceeds Fmax, then the selectivity for that year for age 1 is increased to 0.8 and the fishing mortality recomputed. If this recomputed apical fishing mortality is still above Fmax, the landings are instead limited to those corresponding to Fmax (and this “widened” selectivity). Fmax has been selected as 5% above the maximum that occurred historically. The choice of 0.8 (increased from the 0.6 suggested in Rademeyer and Butterworth (2019)) has been made so as to reduce the chance that the resource is "protected" from undue depletion through inability to make the intended catch rather than by the management rule (MP), and hence provides a more stringent test of the efficacy of that rule.

Step 4: Recruitment

Expected values (in log space) for future recruitments (R_y) are provided by a hockey-stick stock-recruitment relationship:

$$R_y = \begin{cases} R & if\ SSB_y \geq SSB_{threshold} \\ \frac{R}{SSB_{threshold}} SSB_y & if\ SSB_y < SSB_{threshold} \end{cases} \tag{A.6}$$

where

R is the geometric average of the model estimated past (1977-2017) values,

$SSB_{threshold}$ is a fixed value (1.8 million billion eggs produced),

and

$$SSB_y = \sum_{a=2}^{m} f_a N_{y,a} \tag{A.7}$$

with

$f_a = \rho_a mat_a fec_a$ the reproductive output of a female fish of age *a*,

ρ_a is the proportion of female at age *a*,

mat_a is the proportion mature at age *a*, and

fec_a is the fecundity at age *a*.

When projecting, error is added to this expected value, so that for simulation replicate *s*, if

$S = \{\varepsilon_y = lnR_y - lnR: y = 1977, \ldots, 2017\}$, then when projecting:

$R_y^s = Re^{\varepsilon^*}$

where ε^* is drawn at random with replacement from the set I of ε_y values

Although the Recruitment vs Eggs produced plot from the BAM Base Model assessment shows no obvious relationship between the two, clearly there must eventually be some reduction in the number of recruits to be expected as egg production falls. We have taken the fairly standard approach here of assuming a hockey stick relationship whether the hinge-point occurs at the lowest historical annual egg production estimated, though for robustness and precaution a slightly higher value of 180 000 billion eggs was chosen so as to avoid undue influence from the lowest two historical values.

Step 5:

The projected values for numbers-at-age are used to generate values of the abundance indices I_{y+1}^i (in terms of numbers), and similarly for following years. Indices of abundance in future years will not be exactly proportional to true abundance, as they are subject to observation error. Log-normal observation error with autocorrelation is therefore added to the expected value of the abundance index in question (in log space), i.e.:

$$I_y^i = q^i B_y^i e^{\varepsilon_y^i} \tag{A.8}$$

with

$$\varepsilon_y^i = \varphi_y^i - \rho^i \varphi_{y-1}^i \tag{A.9}$$

and φ_y^i from $N\left(0, \left(\sigma^i\right)^2\right)$ (A.10)

where

B_y^i is the abundance available to and indexed by the survey:

$$B_y^i = \sum_{a=1}^{m} S_a^i N_{y,a} e^{-Z_{y,a} T^i/12} \tag{A.11}$$

T^i is the timing of the survey (in month) ($T^i = 6$ for the gill net index and 3 for the seine index).

The autocorrelation coefficient ρ^i for the gillnet index, computed from the historical estimated residuals for the Base Case OM is -0.517 and varies considerably if the relative weighting of the two indices is changed. Negative values of auto-correlation enhance the effective precision of an index, the realism of which is questionable. It was therefore decided to set $\rho^{gill} = 0$ in projections. For the seine index, ρ^i is set at 0.134, the value computed from the historical estimated residuals for the Base Case OM.

The survey selectivities are assumed to remain unchanged. The catchabilities are taken to be those estimated in the OM (the BAM Base Model assessment).

The residual standard deviations σ^i are estimated from the model fit. Since residuals seem to have increased in recent years, the residuals from 2005 onwards have been used for their computation:

$$\sigma^i = \sqrt{\frac{1}{\sum_{y=2005}^{2017} 1} \sum_{y=200}^{2017} \left(lnI_y^i - ln\hat{I}_y^i\right)^2} \tag{A.12}$$

where I_y^i is the observed index value in year y for survey i and $\hat{I}_y^i$ is the corresponding model estimated value this yields σ^i =0.11 for the gill net index and 0.41 for the seine index.

<u>Step 6</u>:

Steps 1-5 are repeated for each future year in turn for as long a period as desired.

Appendix B – The Management Rule (Management Procedure)

The management rule (MP) is empirical. It only overrides and reduces a landing drawn from the historical set if the value of a combined abundance index (see below) falls below a threshold level specified for that index. The basis for the associated computations is set out below:

If $J_y < J_{threshold}$:

$$TAC_{y+1} = \gamma J_y \tag{B.1}$$

where

TAC_y is the catch limit that applies for year *y*,

$J_{threshold}$ (no units) and γ (units: thousand mt) are control parameter (tuning) values (the initial choices (baseline MP) are $J_{threshold} = 0.8$ and $\gamma = 500$); Figure 2 illustrates the rule for these choices for these control parameter values, and

J_y is a measure of the immediate past level in the abundance indices that are available to use for calculations for year *y*:

$$J_y = \frac{1}{p}\sum_{y'=y-p+1}^{y}\left[\left(w_{gill}\frac{I_{y'}^{gill}}{I_{2017}^{gill}} + w_{seine}\frac{I_{y'}^{seine}}{I_{2017}^{seine}}\right)/\left(w_{gill} + w_{seine}\right)\right] \tag{B.2}$$

with

I_y^{gill} and I_y^{seine} being the observed gill net and seine indices, respectively, in year *y*,

w_{gill} and w_{seine} being the weights given to each index ($w_{gill} = 4$ and $w_{seine} = 1$ for the baseline MP, and correspond roughly to inverse variance weighting given the standard deviations of the residuals in the BAM Base Model fit),

and *p* being a control parameter (*p* = 3 for the baseline MP); this parameter is used to smooth away some of the noise in the index by averaging over a few years rather than consider only the most recent year.

Note the assumption has been made that when a TAC is set in year *y* for year *y*+1, values of these abundance indices will be available for the current year *y*.

Appendix C – Performance statistics

Landings-related

1) Average landing 2020-2039
2) Average landing in year where control rule was not applied
3) Average landing in years where control rule was applied
4) Lowest landing over 2020-2039
5) Landing in 2020

Abundance (egg-production, Egg)-related

1) Egg(2020)
2) Egg(2040)
3) Lowest Egg over 2020-2040
4) Probability that Egg(2040) is the lowest in the series from 2020 to 2040 (coarse indication of whether recovery is achieved after a decline)

Catch variability-related

1) Average annual absolute percentage change in landings (AAV) over 2019 to 2039
2) When control rule is applied and landing is decreased, average change in landing

Other

1) Fraction of years from 2020 to 2039 that control rule is applied
2) Fraction of years from 2020 to 2039 that control rule is applied, but actual Egg was above threshold (false positive)
3) Fraction of years from 2020 to 2039 that control rule was not applied, but actual Egg was below threshold (false negative)
4) Fraction of years from 2020 to 2039 that control rule is applied, and actual Egg was below threshold (true positive)
5) Probability that control rule is applied for 2020
6) Fraction of years from 2020 to 2039 that Fmax is hit so that selectivity has to be "spread"
7) Fraction of years from 2020 to 2039 that Fmax is hit and catch cannot be taken despite selectivity being "spread"

Notes:

1) Since for some tests the absolute abundances/egg production will change, so that absolute values might mislead, "rel" statistics are reported which are values relative to the median biomass in the absence of any historical or future catch for that test. For non-stationary situations (e.g. M increasing over time in the future), note that this will change (e.g. the concept of "dynamic B0"), so the "rel" statistic will be relative to this projected value in the year in question.
2) The "threshold" is the lowest historical abundance level, here taken to be in terms of egg production (and denoted SSB in the text).

Appendix D – Base Case and Robustness test OM conditioning results

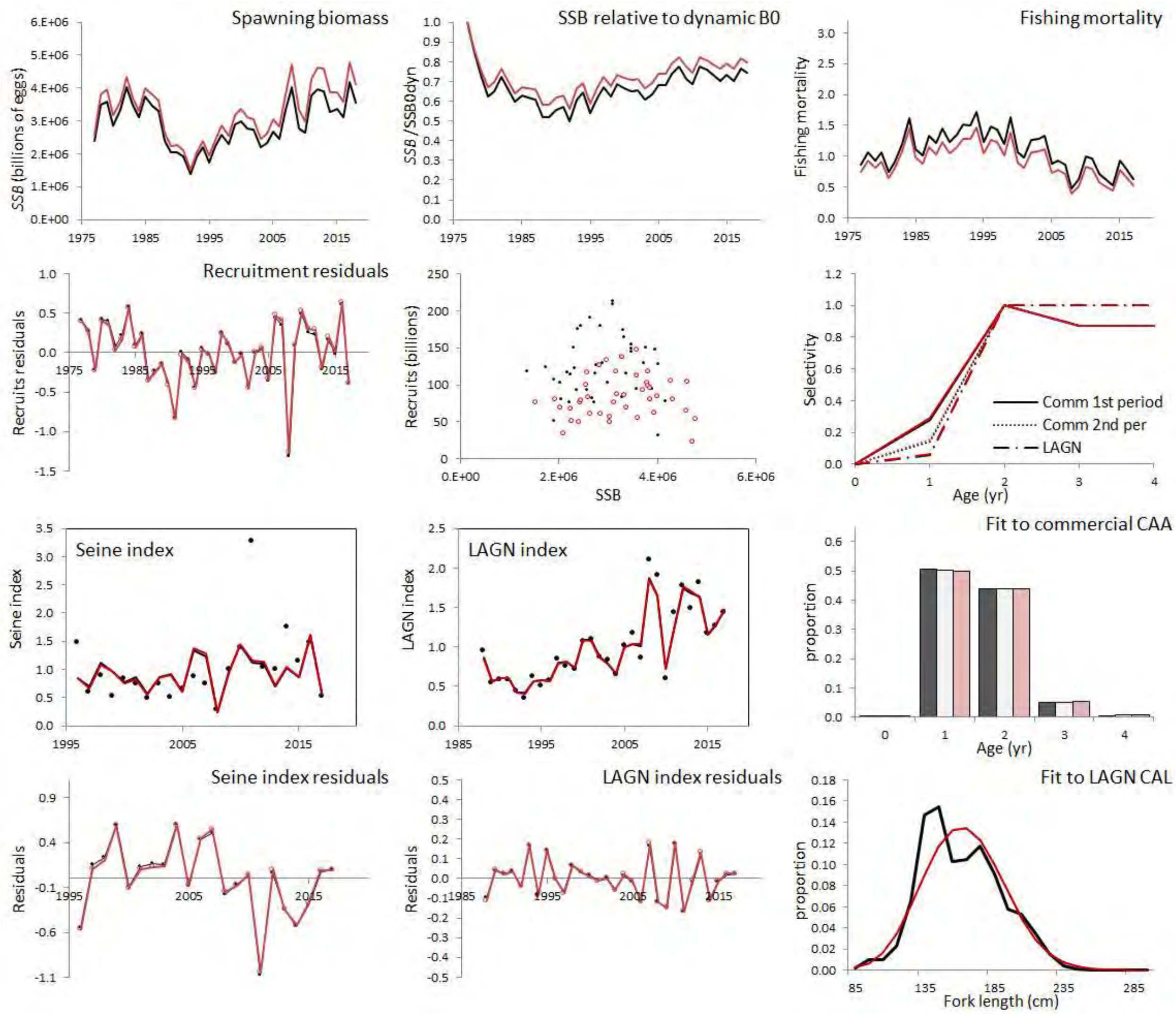


Figure D1a: Assessment results for the Base Case (black lines) and Robustness test 1.1 (M = 1.2) OMs.

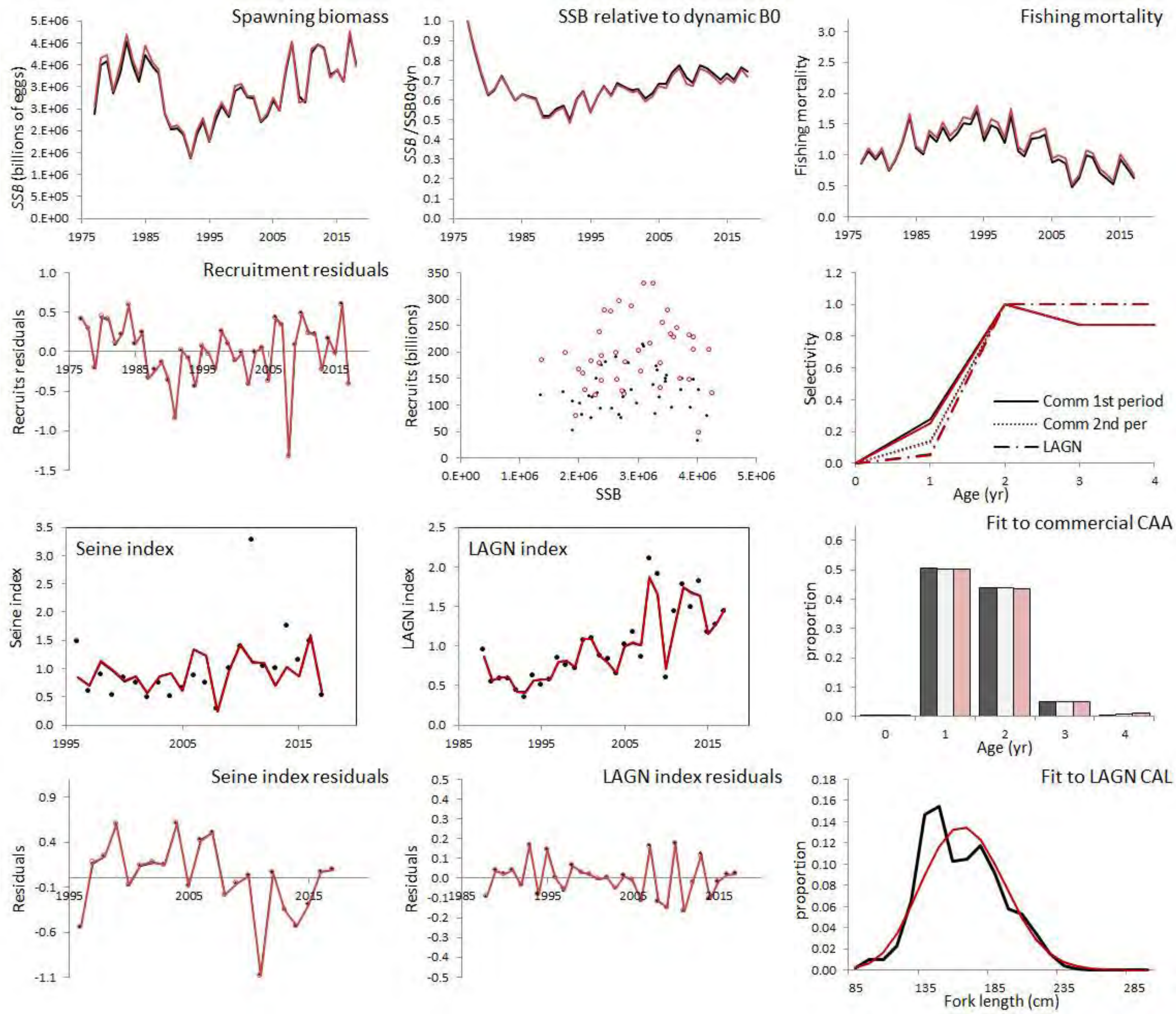


Figure D1b: Assessment results for the Base Case OM (black lines) and Robustness test 1.2 (M'(a)=M(a)*exp(-0.1(a-2))) OMs.

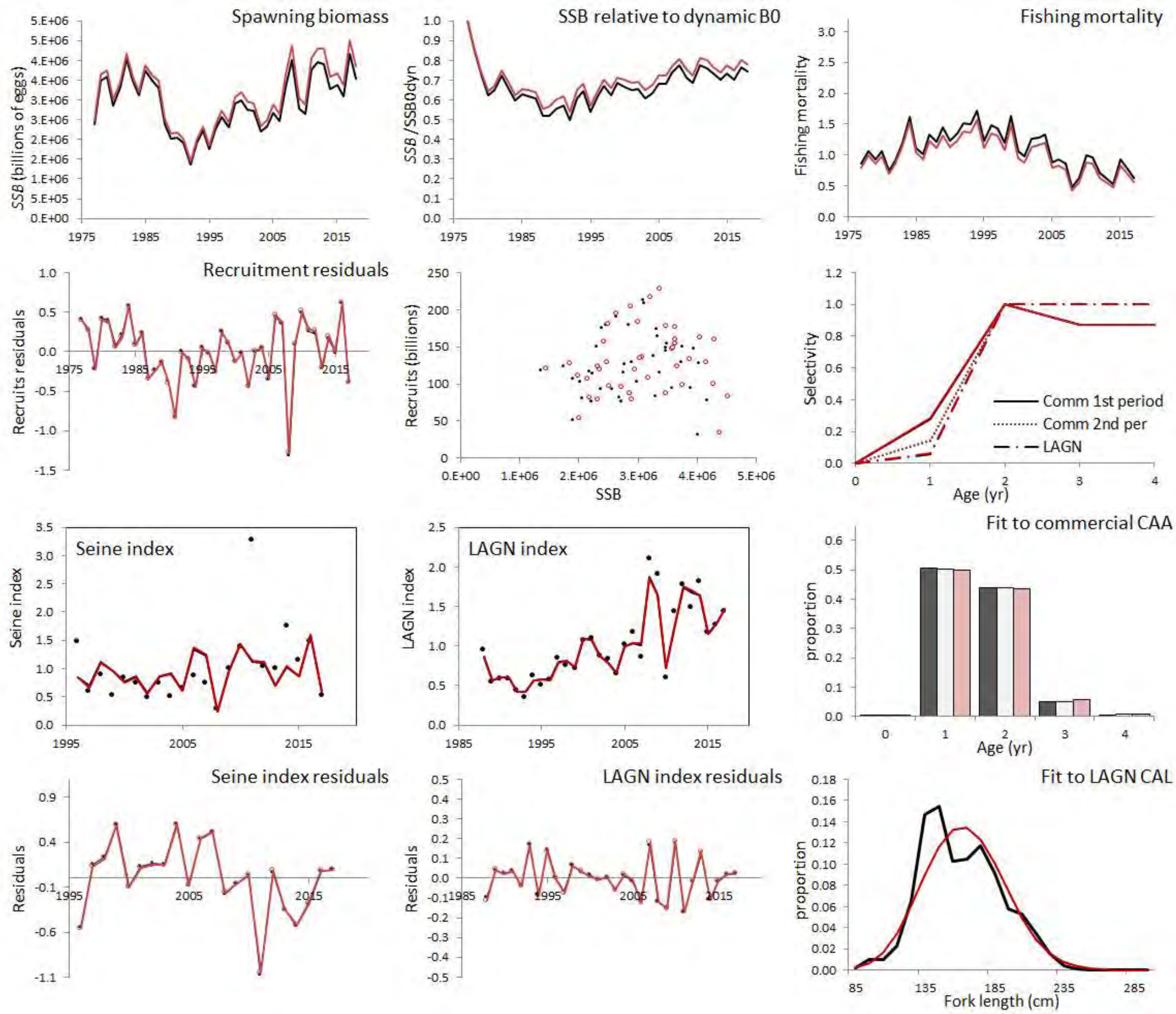


Figure D1c: Assessment results for the Base Case (black lines) and Robustness test 1.3 (M(4+) = 1.67) OMs.

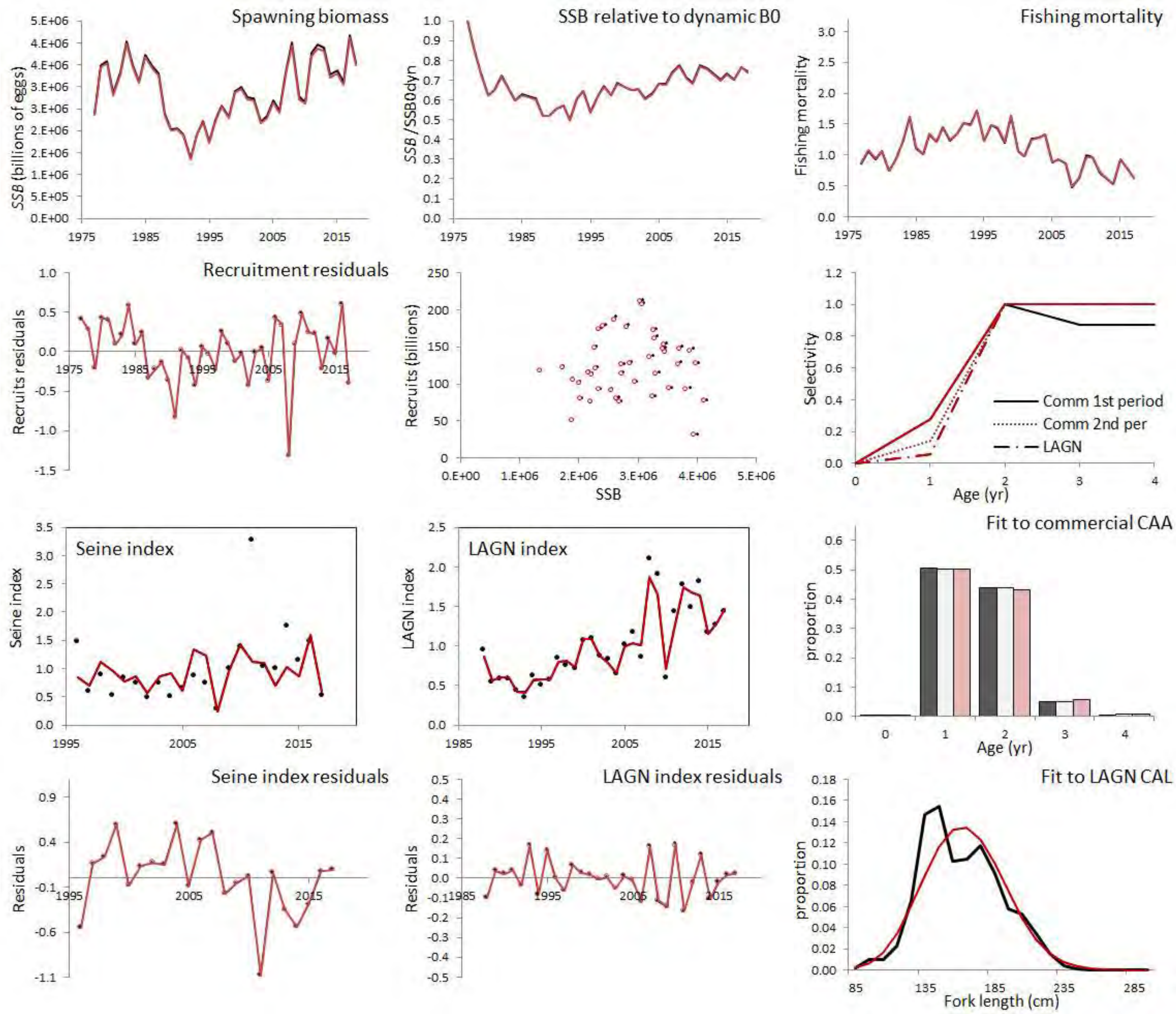


Figure D1d: Assessment results for the Base Case OM (black lines) and Robustness test 2.1 (S(3) = S(4+) = 1.0) OMs.

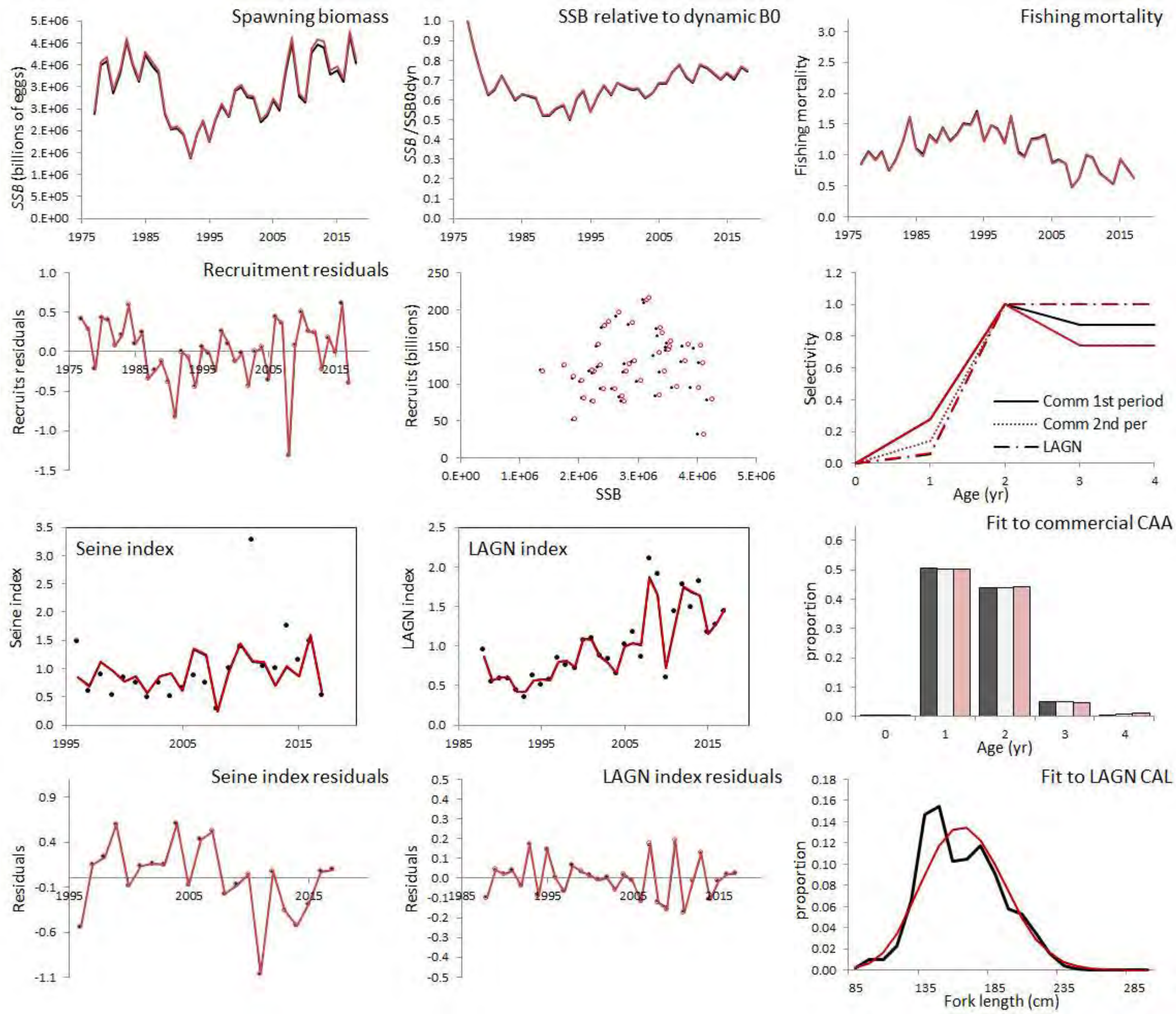


Figure D1e: Assessment results for the Base Case OM (black lines) and Robustness test 2.2 (S(3) = S(4+) = 0.74) OMs.

Base Case (black lines) vs Robustness test 3.1 (I=q*sqrt(B))

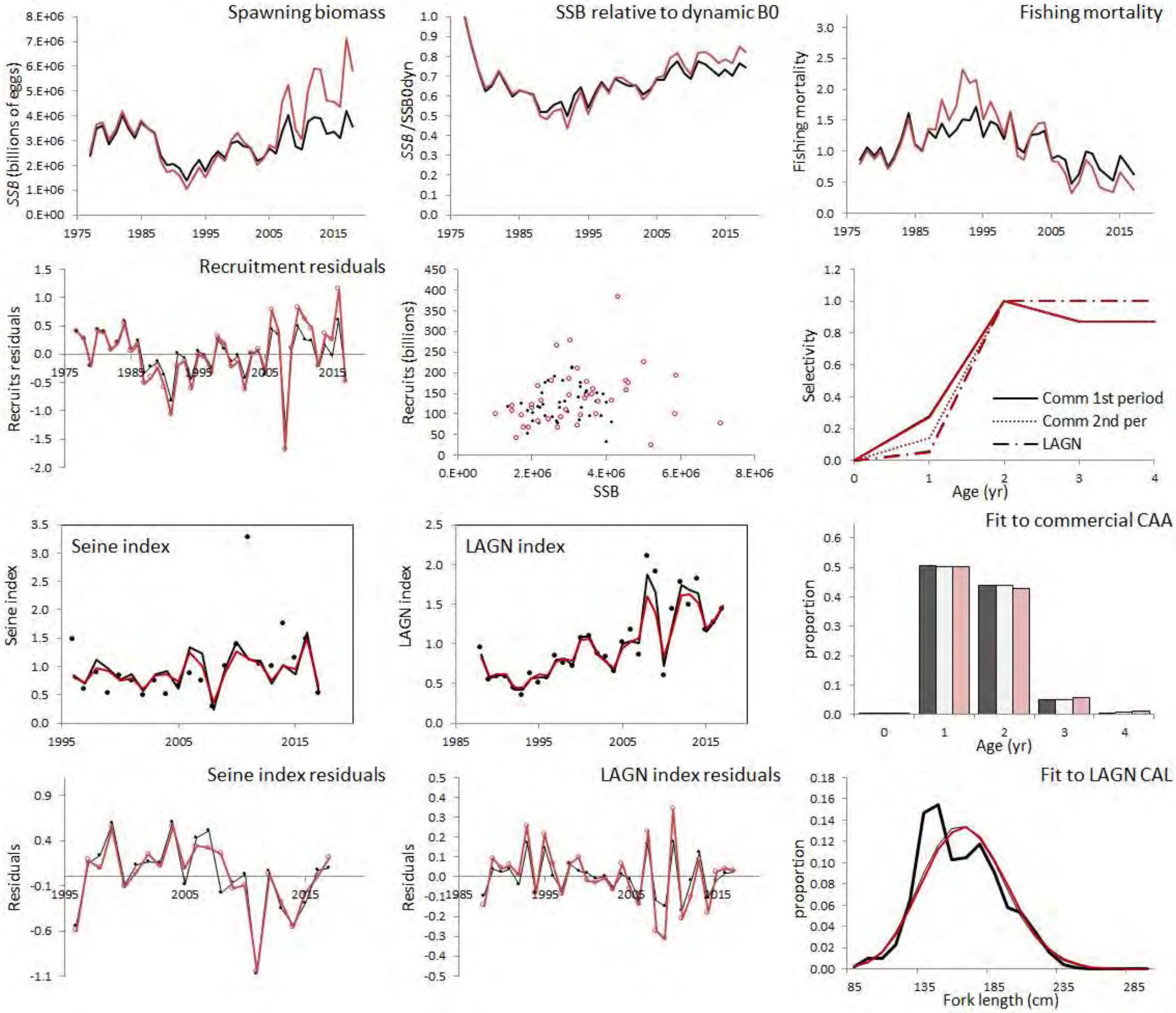


Figure D1f: Assessment results for the Base Case OM (black lines) and Robustness test 3.1 (I=q*sqrt(B)) OMs.

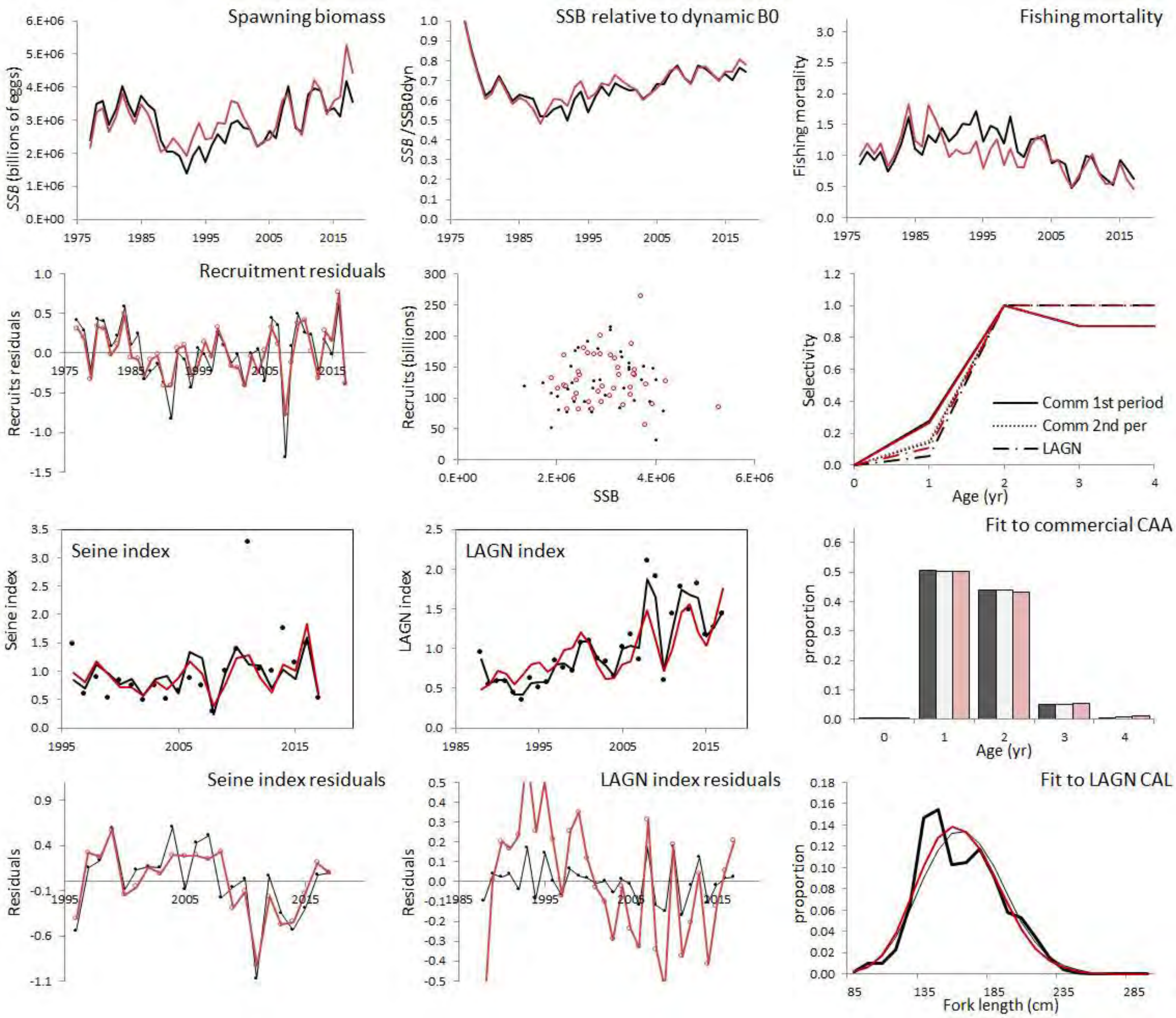


Figure D1g: Assessment results for the Base Case OM (black lines) and Robustness test 3.2 (1:1 weighting) OMs.

Appendix E - Key performance statistics for the Baseline MP under the Base Case OM and all the Robustness tests

Results are shown for the Base Case OM and the different Robustness tests under the baseline MP in Table E1, while Figure E1 plots historical and projected trajectories with and without the baseline MP for all Robustness tests.

Table E1a: Performance statistics for the Base Case OM and Robustness tests 1.1 to 1.5 with and without the management rule (Baseline MP).

Performance_statistics	Base Case			Robustness 1.1			Robustness 1.2			Robustness 1.3			Robustness 1.4			Robustness 1.5		
NO RULE	Median	10	90	Median	10	90	Median	10	90	Median	10	90	Median	10	90	Median	10	90
Related to catch																		
Average landing 2020-2039	494.7	479.2	517.5	494.7	479.2	517.5	494.7	479.2	517.5	494.7	479.2	517.5	363.3	288.9	432.8	484.2	387.8	509.3
Av landing no rule	494.7	479.2	517.5	494.7	479.2	517.5	494.7	479.2	517.5	494.7	479.2	517.5	363.3	288.9	432.8	484.2	387.8	509.3
Av landing with rule	-			-			-			-			-			-		
Lowest landing (2020-2039)	379.9	379.9	425.6	379.9	379.9	425.6	379.9	379.9	425.6	379.9	379.9	425.6	14.6	2.8	69.3	379.9	72.2	400.7
2020 landing	473.7	400.7	590.8	473.7	400.7	590.8	473.7	400.7	590.8	473.7	400.7	590.8	473.7	400.7	590.8	473.7	400.7	590.8
Related to abundance																		
Egg(2020)	3.32	2.73	4.10	3.69	3.05	4.61	3.37	2.72	4.19	3.55	2.91	4.41	3.32	2.73	4.10	3.32	2.73	4.10
Egg(2040)	3.17	2.21	4.04	3.49	2.50	4.53	3.24	2.28	4.17	3.34	2.43	4.25	0.01	0.00	0.07	0.78	0.08	1.97
Egg lowest (2020-2039)	2.00	1.44	2.38	2.17	1.66	2.57	1.99	1.44	2.39	2.10	1.56	2.48	0.01	0.00	0.07	0.76	0.08	1.50
Prob Egg(2040) lowest	6			6			7			5			100			63		
Related to catch variability																		
AAV 2020-2039	0.15	0.12	0.19	0.15	0.12	0.19	0.15	0.12	0.19	0.15	0.12	0.19	0.22	0.17	0.28	0.16	0.12	0.21
AAV with rule	-			-			-			-			-			-		
Other																		
Fraction years rule applied	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
True negative	0.0			0.0			0.0			0.0			0.0			0.0		
Fals negative	0.0			0.0			0.0			0.0			0.0			0.0		
False positive	0.0			0.0			0.0			0.0			0.0			0.0		
True positive	100.0			100.0			100.0			100.0			100.0			100.0		
Prob rule in 2020	0			0			0			0			0			0		
Fraction years Hit Fmax	0	0	0.15	0	0	0.15	0	0	0.15	0	0	0.15	0.6	0.4	0.75	0.35	0.1	0.65
Hit Fmax, landings not taken	0	0	0	0	0	0	0	0	0	0	0	0	0.4	0.25	0.55	0.05	0	0.35
WITH BASELINE MP	Median	10	90	Median	10	90	Median	10	90	Median	10	90	Median	10	90	Median	10	90
Related to catch																		
Average landing 2020-2039	483.8	448.6	511.0	481.5	448.6	510.8	484.2	448.6	511.0	482.8	449.3	511.0	333.7	285.0	401.2	423.9	368.9	466.1
Av landing no rule	494.8	477.8	519.1	493.7	476.0	517.5	494.8	477.2	519.1	494.5	476.3	519.1	501.4	471.7	526.9	497.8	477.0	519.6
Av landing with rule	344.4	0.0	389.5	353.7	0.0	387.1	338.5	0.0	389.7	354.4	0.0	390.4	226.4	136.5	292.6	321.0	234.5	360.0
Lowest landing (2020-2039)	364.2	272.8	400.7	357.3	276.7	395.6	368.5	273.9	400.7	354.1	272.0	396.1	75.4	19.9	185.9	246.9	138.4	336.1
2020 landing	473.7	400.7	590.8	473.7	400.7	590.8	473.7	400.7	590.8	473.7	400.7	590.8	473.7	400.7	590.8	473.7	400.7	590.8
Related to abundance																		
Egg(2020)	3.32	2.73	4.10	3.69	3.05	4.61	3.37	2.72	4.19	3.55	2.91	4.41	3.32	2.73	4.10	3.32	2.73	4.10
Egg(2040)	3.19	2.37	4.04	3.54	2.56	4.53	3.26	2.41	4.17	3.38	2.47	4.25	0.17	0.04	0.51	1.59	0.87	2.33
Egg lowest (2020-2039)	2.01	1.48	2.38	2.21	1.69	2.57	2.02	1.52	2.40	2.12	1.62	2.48	0.17	0.04	0.51	1.23	0.67	1.66
Prob Egg(2040) lowest	5			5			6			4			96			25		
Related to catch variability																		
AAV 2020-2039	0.16	0.13	0.20	0.17	0.13	0.21	0.16	0.13	0.20	0.17	0.13	0.21	0.18	0.13	0.23	0.17	0.12	0.23
AAV with rule	0.15	0.00	0.26	0.16	0.00	0.27	0.15	0.00	0.26	0.16	0.00	0.29	0.18	0.13	0.27	0.17	0.10	0.24
LAgill																		
Fraction years rule applied	0.10	0.00	0.25	0.10	0.00	0.25	0.10	0.00	0.25	0.10	0.00	0.25	0.60	0.45	0.70	0.40	0.20	0.60
True negative	10.1			11.5			9.4			10.9			63.2			43.7		
Fals negative	0.8			1.2			0.9			1.0			0.3			1.4		
False positive	6.8			7.6			6.8			7.6			6.9			12.5		
True positive	82.4			79.8			83.0			80.6			29.7			42.5		
Prob rule in 2020	0			0			0			0			0			0		
Fraction years Hit Fmax	0	0	0.1	0	0	0.05	0	0	0.1	0	0	0.1	0.3	0.15	0.45	0.1	0	0.2
Hit Fmax, landings not taken	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.05	0	0	0

Table E1b: Performance statistics for the Base Case OM and Robustness tests 2.1 to 2.3 with and without the management rule (Baseline MP).

Performance_statistics	Base Case			Robustness 2.1			Robustness 2.2			Robustness 2.3		
NO RULE	Median	10	90	Median	10	90	Median	10	90	Median	10	90
Related to catch												
Average landing 2020-2039	494.7	479.2	517.5	494.7	479.2	517.5	494.7	479.2	517.5	494.7	479.2	517.5
Av landing no rule	494.7	479.2	517.5	494.7	479.2	517.5	494.7	479.2	517.5	494.7	479.2	517.5
Av landing with rule	-			-			-			-		
Lowest landing (2020-2039)	379.9	379.9	425.6	379.9	379.9	425.6	379.9	379.9	425.6	379.9	379.9	425.6
2020 landing	473.7	400.7	590.8	473.7	400.7	590.8	473.7	400.7	590.8	473.7	400.7	590.8
Related to abundance												
Egg(2020)	3.32	2.73	4.10	3.27	2.70	4.02	3.39	2.78	4.22	3.34	2.75	4.12
Egg(2040)	3.17	2.21	4.04	3.13	2.14	4.00	3.24	2.27	4.11	3.18	2.21	4.05
Egg lowest (2020-2039)	2.00	1.44	2.38	1.97	1.37	2.36	2.02	1.48	2.41	1.99	1.40	2.38
Prob Egg(2040) lowest	6			6			5			6		
Related to catch variability												
AAV 2020-2039	0.15	0.12	0.19	0.15	0.12	0.19	0.15	0.12	0.19	0.15	0.12	0.19
AAV with rule	-			-			-			-		
Other												
Fraction years rule applied	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
True negative	0.0			0.0			0.0			0.0		
Fals negative	0.0			0.0			0.0			0.0		
False positive	0.0			0.0			0.0			0.0		
True positive	100.0			100.0			100.0			100.0		
Prob rule in 2020	0			0			0			0		
Fraction years Hit Fmax	0	0	0.15	0	0	0.15	0	0	0.15	0	0	0.05
Hit Fmax, landings not taken	0	0	0	0	0	0	0	0	0	0	0	0
WITH BASELINE MP	Median	10	90	Median	10	90	Median	10	90	Median	10	90
Related to catch												
Average landing 2020-2039	483.8	448.6	511.0	484.0	448.6	511.0	483.8	448.6	511.0	486.4	448.6	512.0
Av landing no rule	494.8	477.8	519.1	494.8	477.8	519.1	494.8	476.3	519.1	493.8	478.1	518.5
Av landing with rule	344.4	0.0	389.5	346.8	0.0	392.5	346.5	0.0	388.2	315.5	0.0	381.2
Lowest landing (2020-2039)	364.2	272.8	400.7	365.2	273.5	400.7	362.2	272.6	400.7	374.7	274.2	400.7
2020 landing	473.7	400.7	590.8	473.7	400.7	590.8	473.7	400.7	590.8	473.7	400.7	590.8
Related to abundance												
Egg(2020)	3.32	2.73	4.10	3.27	2.70	4.02	3.39	2.78	4.22	3.34	2.75	4.12
Egg(2040)	3.19	2.37	4.04	3.13	2.34	4.00	3.26	2.40	4.11	3.18	2.34	4.05
Egg lowest (2020-2039)	2.01	1.48	2.38	1.98	1.46	2.36	2.05	1.50	2.41	2.00	1.45	2.38
Prob Egg(2040) lowest	5			5			4			5		
Related to catch variability												
AAV 2020-2039	0.16	0.13	0.20	0.16	0.13	0.20	0.17	0.13	0.20	0.16	0.13	0.20
AAV with rule	0.15	0.00	0.26	0.15	0.00	0.27	0.15	0.00	0.27	0.12	0.00	0.25
LAgill												
Fraction years rule applied	0.10	0.00	0.25	0.10	0.00	0.25	0.10	0.00	0.25	0.05	0.00	0.25
True negative	10.1			9.9			10.2			8.9		
Fals negative	0.8			0.9			0.8			0.6		
False positive	6.8			6.7			7.0			5.7		
True positive	82.4			82.7			82.1			84.9		
Prob rule in 2020	0			0			0			0		
Fraction years Hit Fmax	0	0	0.1	0	0	0.1	0	0	0.1	0	0	0
Hit Fmax, landings not taken	0	0	0	0	0	0	0	0	0	0	0	0

Table E1c: Performance statistics for the Base Case OM and Robustness tests 3.1 to 3.6 with and without the management rule (Baseline MP).

Performance_statistics	Base Case			Robustness 3.1			Robustness 3.2			Robustness 3.3			Robustness 3.4			Robustness 3.5			Robustness 3.6		
NO RULE	Median	10	90	Median	10	90	Median	10	90	Median	10	90	Median	10	90	Median	10	90	Median	10	90
Related to catch																					
Average landing 2020-2039	494.7	479.2	517.5	494.4	477.2	517.5	494.7	480.0	517.5	494.7	479.2	517.5	494.7	479.2	517.5	494.7	479.2	517.5	494.7	479.2	517.5
Av landing no rule	494.7	479.2	517.5	494.4	477.2	517.5	494.7	480.0	517.5	494.7	479.2	517.5	494.7	479.2	517.5	494.7	479.2	517.5	494.7	479.2	517.5
Av landing with rule	-			-			-			-			-			-			-		
Lowest landing (2020-2039)	379.9	379.9	425.6	379.9	379.9	425.6	379.9	379.9	425.6	379.9	379.9	425.6	379.9	379.9	425.6	379.9	379.9	425.6	379.9	379.9	425.6
2020 landing	473.7	400.7	590.8	473.7	400.7	590.8	473.7	400.7	590.8	473.7	400.7	590.8	473.7	400.7	590.8	473.7	400.7	590.8	473.7	400.7	590.8
Related to abundance																					
Egg(2020)	3.32	2.73	4.10	4.15	3.30	5.82	3.63	3.14	4.31	3.32	2.73	4.10	3.32	2.73	4.10	3.32	2.73	4.10	3.32	2.73	4.10
Egg(2040)	3.17	2.21	4.04	3.40	2.17	4.89	3.29	2.56	4.03	3.17	2.21	4.04	3.17	2.21	4.04	3.17	2.21	4.04	3.17	2.21	4.04
Egg lowest (2020-2039)	2.00	1.44	2.38	1.90	1.12	2.43	2.26	1.88	2.63	2.00	1.44	2.38	2.00	1.44	2.38	2.00	1.44	2.38	2.00	1.44	2.38
Prob Egg(2040) lowest	6			4			5			6			6			6			6		
Related to catch variability																					
AAV 2020-2039	0.15	0.12	0.19	0.15	0.12	0.19	0.15	0.12	0.19	0.15	0.12	0.19	0.15	0.12	0.19	0.15	0.12	0.19	0.15	0.12	0.19
AAV with rule	-			-			-			-			-			-			-		
Other																					
Fraction years rule applied	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
True negative	0.0			0.0			0.0			0.0			0.0			0.0			0.0		
Fals negative	0.0			0.0			0.0			0.0			0.0			0.0			0.0		
False positive	0.0			0.0			0.0			0.0			0.0			0.0			0.0		
True positive	100.0			100.0			100.0			100.0			100.0			100.0			100.0		
Prob rule in 2020	0			0			0			0			0			0			0		
Fraction years Hit Fmax	0	0	0.15	0	0	0.1	0	0	0	0	0	0.15	0	0	0.15	0	0	0.15	0	0	0.15
Hit Fmax, landings not taken	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
WITH BASELINE MP	Median	10	90	Median	10	90	Median	10	90	Median	10	90	Median	10	90	Median	10	90	Median	10	90
Related to catch																					
Average landing 2020-2039	483.8	448.6	511.0	494.4	477.2	517.5	494.0	475.0	516.0	483.2	448.4	508.2	482.2	448.1	508.1	481.9	447.1	507.4	483.1	448.6	511.0
Av landing no rule	494.8	477.8	519.1	494.4	477.2	517.5	495.0	479.2	518.0	493.7	476.3	519.1	493.8	477.2	517.6	494.0	478.1	517.8	494.5	478.1	519.1
Av landing with rule	344.4	0.0	389.5	0.0	0.0	0.0	0.0	0.0	383.8	340.0	0.0	388.1	343.1	0.0	389.7	337.7	0.0	385.8	345.4	0.0	389.1
Lowest landing (2020-2039)	364.2	272.8	400.7	379.9	379.9	425.6	379.9	379.9	400.7	367.8	270.9	400.7	357.6	267.3	400.7	353.9	256.7	400.7	359.0	270.4	396.5
2020 landing	473.7	400.7	590.8	473.7	400.7	590.8	473.7	400.7	590.8	473.7	400.7	590.8	473.7	400.7	590.8	473.7	400.7	590.8	473.7	400.7	590.8
Related to abundance																					
Egg(2020)	3.32	2.73	4.10	4.15	3.30	5.82	3.63	3.14	4.31	3.32	2.73	4.10	3.32	2.73	4.10	3.32	2.73	4.10	3.32	2.73	4.10
Egg(2040)	3.19	2.37	4.04	3.40	2.17	4.89	3.29	2.58	4.03	3.19	2.37	4.04	3.20	2.35	4.05	3.20	2.35	4.05	3.19	2.37	4.04
Egg lowest (2020-2039)	2.01	1.48	2.38	1.90	1.12	2.43	2.26	1.88	2.63	2.01	1.48	2.39	2.01	1.48	2.39	2.01	1.48	2.39	2.01	1.51	2.38
Prob Egg(2040) lowest	5			4			3			5			5			5			5		
Related to catch variability																					
AAV 2020-2039	0.16	0.13	0.20	0.15	0.12	0.19	0.15	0.12	0.19	0.16	0.12	0.21	0.17	0.13	0.21	0.17	0.13	0.22	0.17	0.13	0.21
AAV with rule	0.15	0.00	0.26	0.00	0.00	0.00	0.00	0.00	0.21	0.15	0.00	0.31	0.14	0.00	0.29	0.17	0.00	0.30	0.15	0.00	0.28
LAgill																					
Fraction years rule applied	0.10	0.00	0.25	0.00	0.00	0.00	0.00	0.00	0.05	0.10	0.00	0.25	0.10	0.00	0.25	0.10	0.00	0.25	0.10	0.00	0.25
True negative	10.1			0.1			0.9			9.8			9.4			8.4			10.7		
Fals negative	0.8			0.0			0.4			1.3			2.3			3.9			0.9		
False positive	6.8			0.1			3.1			7.0			7.5			8.3			7.3		
True positive	82.4			99.9			95.6			82.0			81.0			79.5			81.2		
Prob rule in 2020	0			0			0			0			0			0			0		
Fraction years Hit Fmax	0	0	0.1	0	0	0.1	0	0	0	0	0	0.1	0	0	0.1	0	0	0.1	0	0	0.1
Hit Fmax, landings not taken	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Table E1d: Performance statistics for the Base Case OM and Robustness tests 4.1 to 7.1 with and without the management rule (Baseline MP).

Performance_statistics	Base Case			Robustness 4.1			Robustness 5.1			Robustness 6.1			Robustness 7.1		
NO RULE	Median	10	90	Median	10	90	Median	10	90	Median	10	90	Median	10	90
Related to catch															
Average landing 2020-2039	494.7	479.2	517.5	371.9	169.9	500.7	493.7	464.2	516.0	494.0	477.2	517.4	494.7	479.2	517.5
Av landing no rule	494.7	479.2	517.5	371.9	169.9	500.7	493.7	464.2	516.0	494.0	477.2	517.4	494.7	479.2	517.5
Av landing with rule	-			-			-			-			-		
Lowest landing (2020-2039)	379.9	379.9	425.6	181.9	41.0	379.9	379.9	379.9	400.7	379.9	379.9	425.6	379.9	379.9	425.6
2020 landing	473.7	400.7	590.8	473.7	400.7	590.8	473.7	400.7	590.8	473.7	400.7	590.8	473.7	400.7	590.8
Related to abundance															
Egg(2020)	3.32	2.73	4.10	3.32	2.73	4.10	3.32	2.73	4.10	3.23	2.64	4.01	3.32	2.73	4.10
Egg(2040)	3.17	2.21	4.04	0.59	0.10	3.80	2.87	0.84	3.90	3.05	2.03	3.91	3.17	2.21	4.04
Egg lowest (2020-2039)	2.00	1.44	2.38	0.30	0.05	1.06	1.98	0.80	2.38	1.89	1.16	2.26	2.00	1.44	2.38
Prob Egg(2040) lowest	6			9			12			6			6		
Related to catch variability															
AAV 2020-2039	0.15	0.12	0.19	0.20	0.14	0.27	0.15	0.12	0.20	0.15	0.12	0.19	0.15	0.12	0.19
AAV with rule	-			-			-			-			-		
Other															
Fraction years rule applied	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
True negative	0.0			0.0			0.0			0.0			0.0		
Fals negative	0.0			0.0			0.0			0.0			0.0		
False positive	0.0			0.0			0.0			0.0			0.0		
True positive	100.0			100.0			100.0			100.0			100.0		
Prob rule in 2020	0			0			0			0			0		
Fraction years Hit Fmax	0	0	0.15	0.85	0.25	0.9	0.025	0	0.35	0.1	0	0.3	0	0	0.15
Hit Fmax, landings not taken	0	0	0	0.65	0	0.85	0	0	0.05	0	0	0	0	0	0
WITH BASELINE MP	Median	10	90	Median	10	90	Median	10	90	Median	10	90	Median	10	90
Related to catch															
Average landing 2020-2039	488.7	454.7	512.6	382.4	204.5	459.3	484.9	435.7	513.6	482.4	444.9	513.1	488.7	454.7	513.6
Av landing no rule	494.8	476.3	518.5	498.1	467.3	529.6	495.4	476.0	521.1	495.8	476.0	520.3	494.8	476.3	520.4
Av landing with rule	nan	nan	498.1	277.2	145.9	389.0	404.3	384.0	498.1	402.7	383.0	475.1	nan	nan	498.1
Lowest landing (2020-2039)	379.9	312.1	400.7	176.5	43.3	300.1	379.9	265.1	400.7	379.9	303.4	400.7	379.9	312.1	400.7
2020 landing	473.7	400.7	590.8	473.7	400.7	590.8	473.7	400.7	590.8	473.7	400.7	590.8	473.7	400.7	590.8
Related to abundance															
Egg(2020)	3.32	2.73	4.10	3.32	2.73	4.10	3.32	2.73	4.10	3.23	2.64	4.01	3.32	2.73	4.10
Egg(2040)	3.18	2.36	4.04	3.16	2.18	4.01	3.15	1.96	4.00	3.09	2.26	3.95	3.18	2.36	4.04
Egg lowest (2020-2039)	2.01	1.47	2.38	0.51	0.14	1.31	1.99	1.07	2.38	1.90	1.27	2.28	2.01	1.47	2.38
Prob Egg(2040) lowest	6			0			7			6			6		
Related to catch variability															
AAV 2020-2039	0.15	0.11	0.19	0.18	0.14	0.24	0.15	0.12	0.19	0.15	0.11	0.19	0.15	0.11	0.19
AAV with rule	0.10	0.08	nan	0.16	0.10	0.27	0.10	0.08	nan	0.10	0.07	nan	0.10	nan	0.10
LAgill															
Fraction years rule applied	0.10	0.00	0.30	0.53	0.35	0.80	0.10	0.00	0.40	0.15	0.00	0.35	0.10	0.00	0.30
True negative	10.8			55.5			15.5			16.0			11.0		
Fals negative	0.8			0.4			0.9			1.0			0.8		
False positive	7.4			5.8			6.9			8.2			7.4		
True positive	81.0			38.3			76.8			74.9			80.9		
Prob rule in 2020	0			0			0			0			0		
Fraction years Hit Fmax	0	0	0.1	0.3	0.1	0.45	0	0	0.2	0.05	0	0.2	0	0	0.1
Hit Fmax, landings not taken	0	0	0	0.05	0	0.15	0	0	0	0	0	0	0	0	0

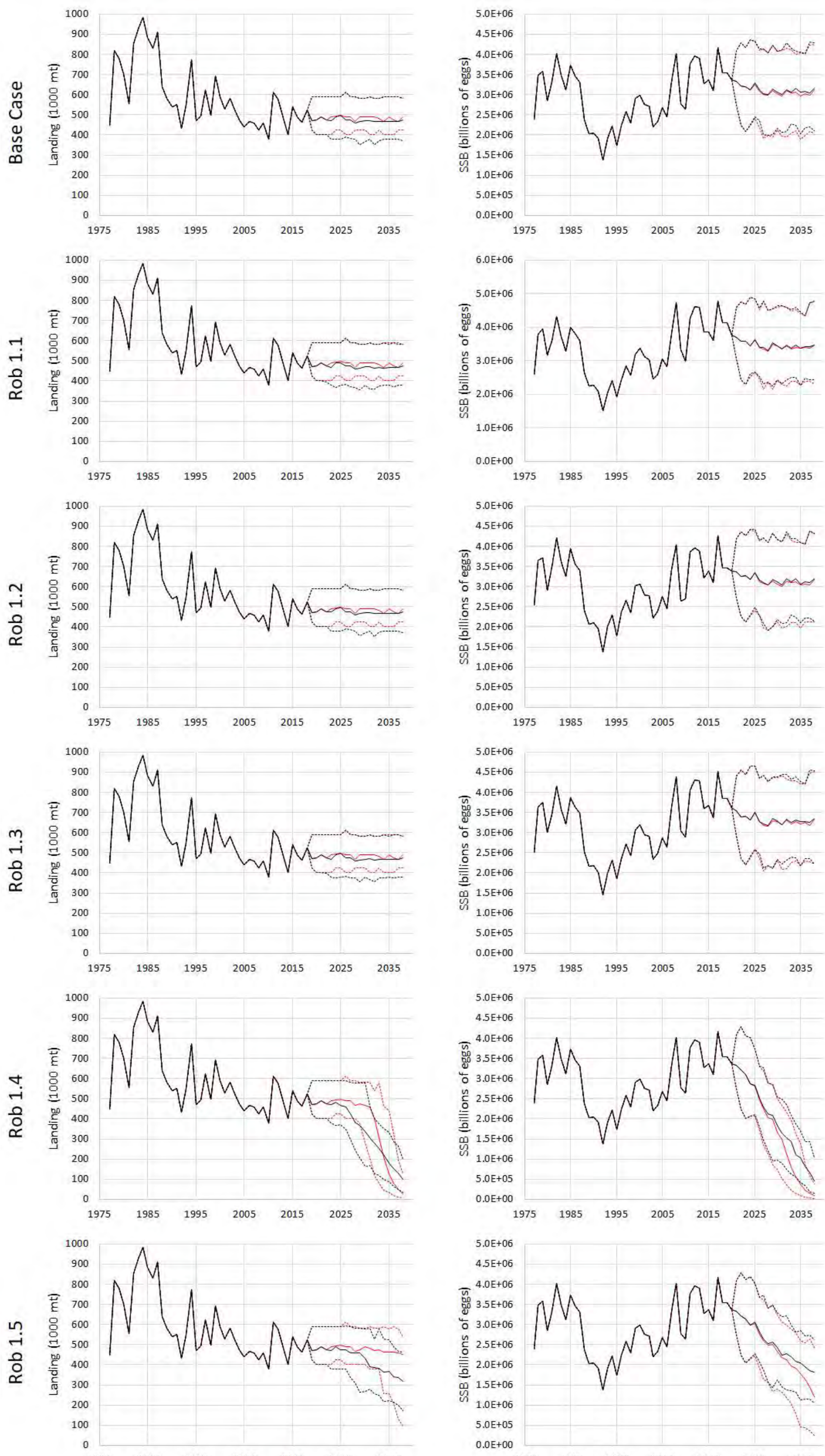


Fig E1a: Historical estimates and projected 20-year median with 10%- and 90%-iles landing and SSB without (red lines) and with the management rule (Baseline MP, black lines) for the Base Case OM and **Robustness test 1.1 to 1.5** OMs.

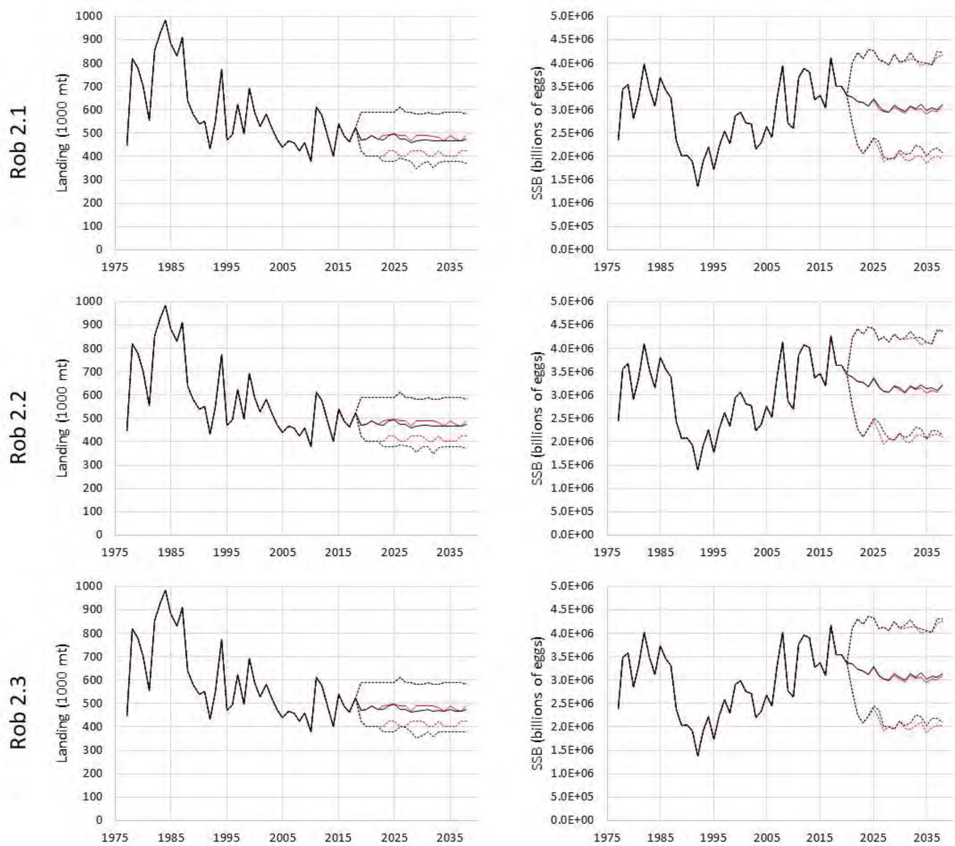


Fig E1b: Historical estimates and projected 20-year median with 10%- and 90%-iles landing and SSB without (red lines) and with the management rule (Baseline MP, black lines) for **Robustness test 2.1 to 2.3** OMs.

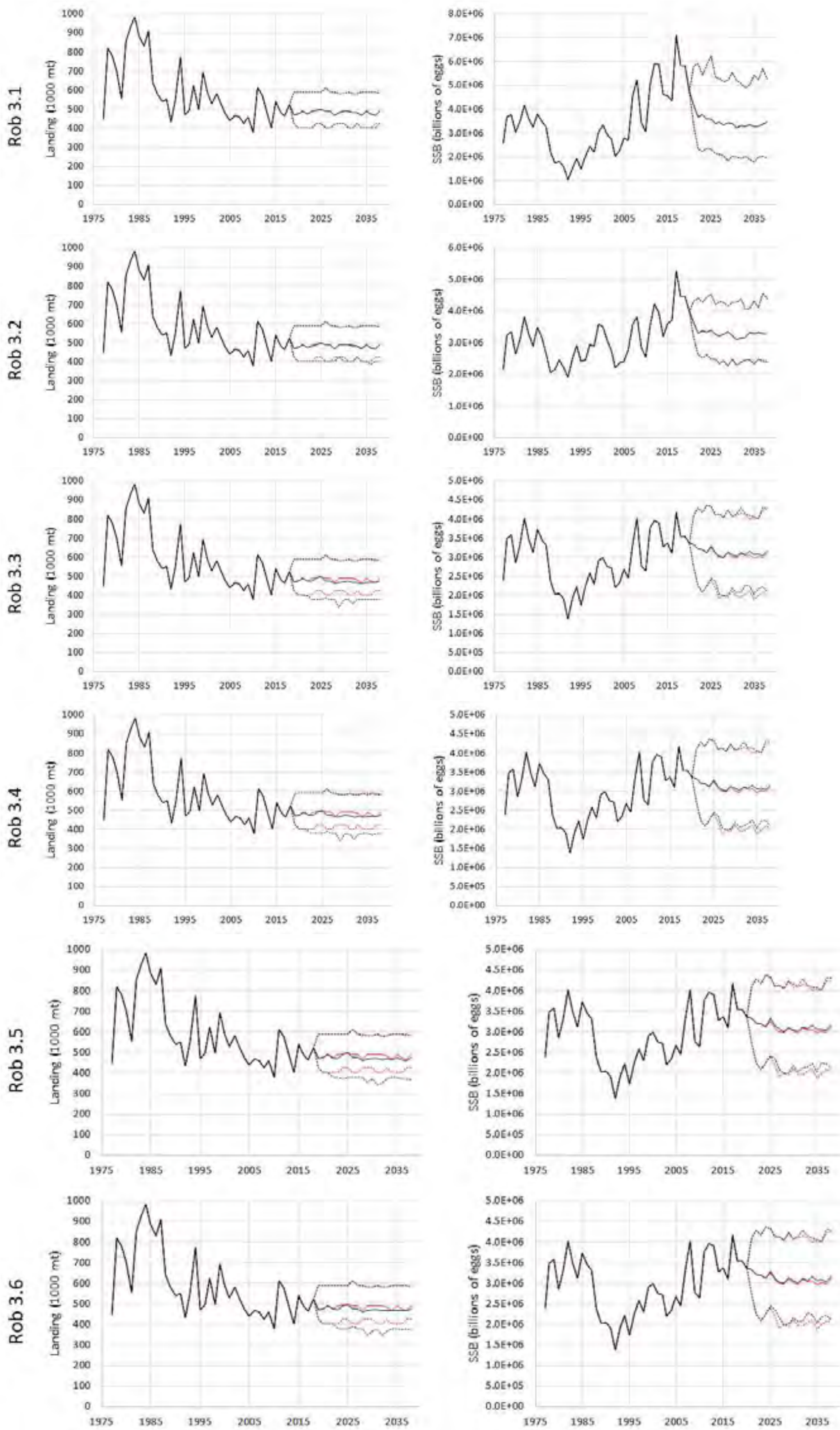


Fig E1c: Historical estimates and projected 20-year median with 10%- and 90%-iles landing and SSB without (red lines) and with the management rule (Baseline MP, black lines) for **robustness test 3.1 to 3.6** OMs.

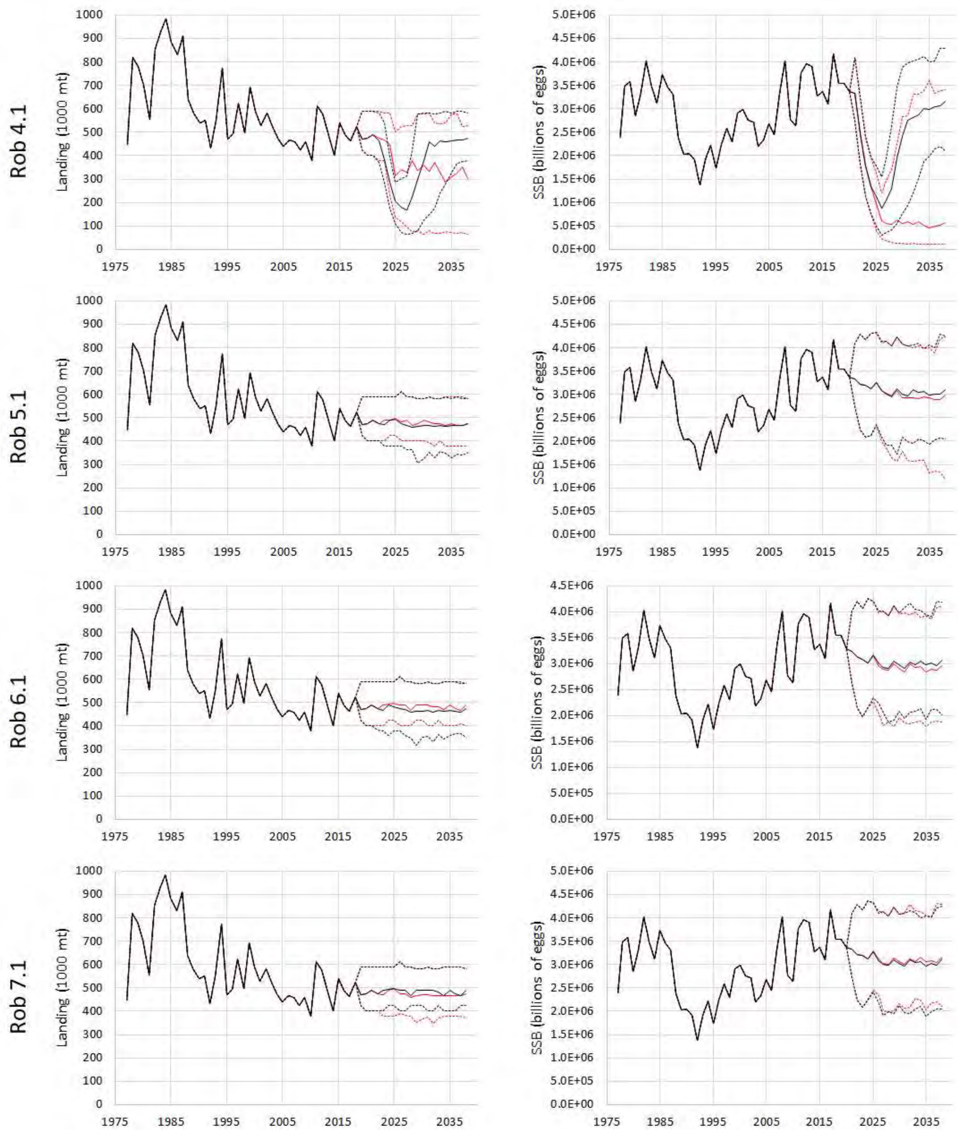


Fig E1d: Historical estimates and projected 20-year median with 10%- and 90%-iles landing and SSB without (red lines) and with the management rule (Baseline MP, black lines) for **Robustness test 4.1 to 7.1** OMs.

Appendix F - Key performance statistics when varying the control parameters of the Baseline MP

Results are shown for varying each of the control parameter values of the baseline MP in turn, and compared to the results in the absence of any control which is shown by the red dot. Medians with 10%- and 90%-iles are shown. Results for the value used in the Baseline MP are indicated by a triangle. The red line shown for the "SSB lowest" plots is the lowest value historically as estimated for the Base Case OM. The plots cover the Base Case OM and the two important Robustness tests: 1.5 (increasing *M* in the future)) and 4.1 (poor recruitment in the future).

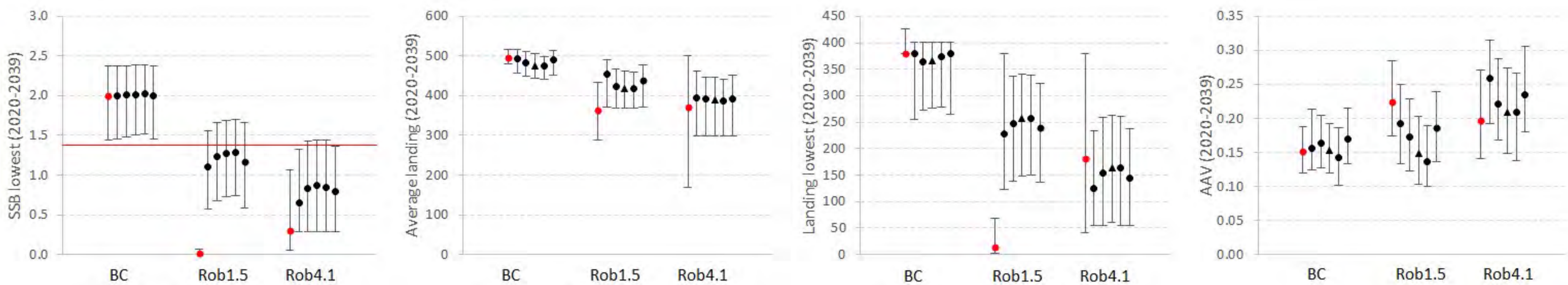


Figure F1a: Performance statistics for no harvest control rule and for **varying values of the $J_{threshold}$** control parameter: 0.60, 0.7, 0.80 (Baseline MP), 0.9 and 1.0. Results are shown for the Base Case OM with the 1.5 and 4.1 Robustness tests.

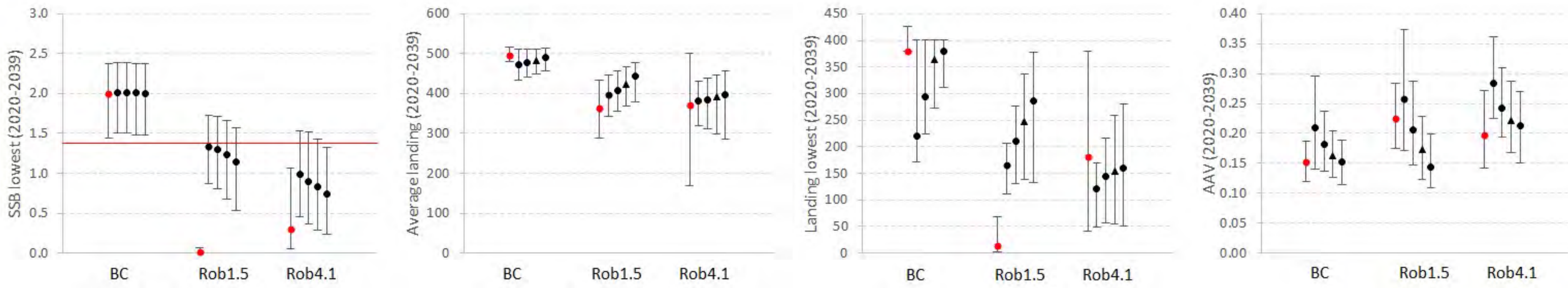


Figure F1b: Performance statistics for no harvest control rule and for **varying values of the** γ control parameter: 300, 400, 500 (Baseline MP), and 600. Results are shown for the Base Case OM with the 1.5 and 4.1 Robustness tests.

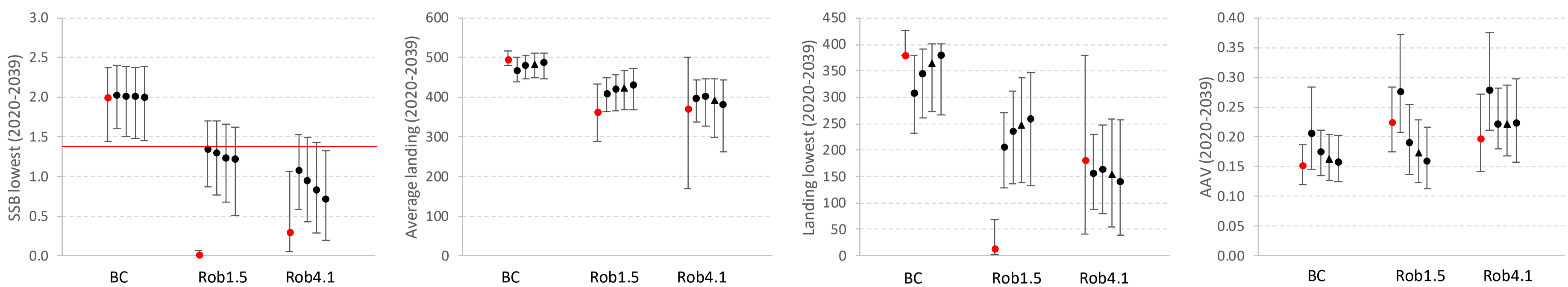


Figure F1c: Performance statistics for no harvest control rule and for **varying values of the *p*** control parameter: 2, 3 (Baseline MP) and 4. Results are shown for the Base Case OM with the 1.5 and 4.1 Robustness tests.

Appendix B - Further resultsC_18 July

Thursday 18 July

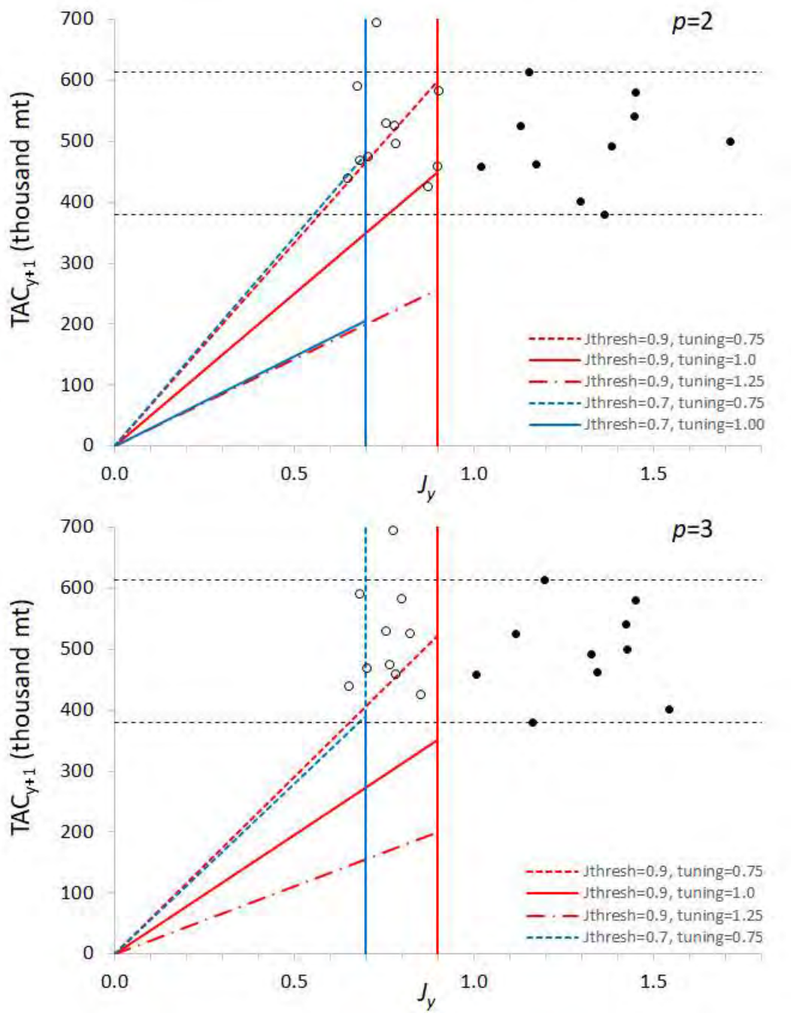


Figure 1: Illustration of the MP variants' rule for two value of the $J_{threshold}$ control parameter, with the γ control parameter value tuned so that the median lowest SSB for Robustness test 4.1 (t) is equal to 0.75, 1.0 and 1.25. In some cases, the target could not be reached (Jthresh=0.7, t=1.25 for p=2 and Jthresh=0.7, t=1.00 and 1.25 for p=3) The horizontal dash lines show the 2009-2018 minimum and maximum landing values. The last 10 years' historical (2008-2017) J_y vs TAC_{y+1} are shown as black dots.

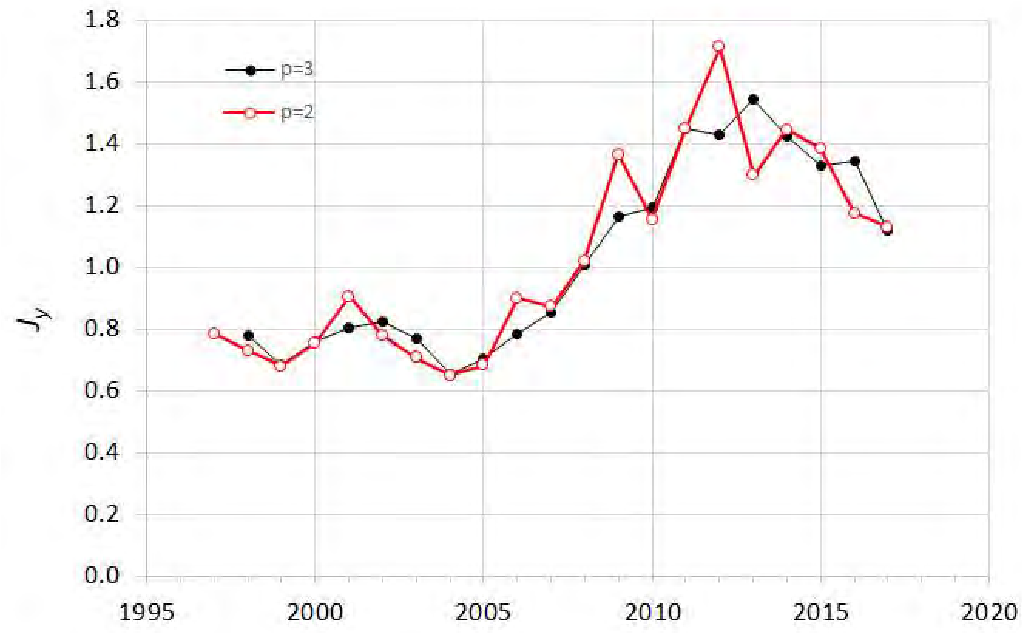


Figure 2: Historical combined index J_y values for p=2 and p=3.

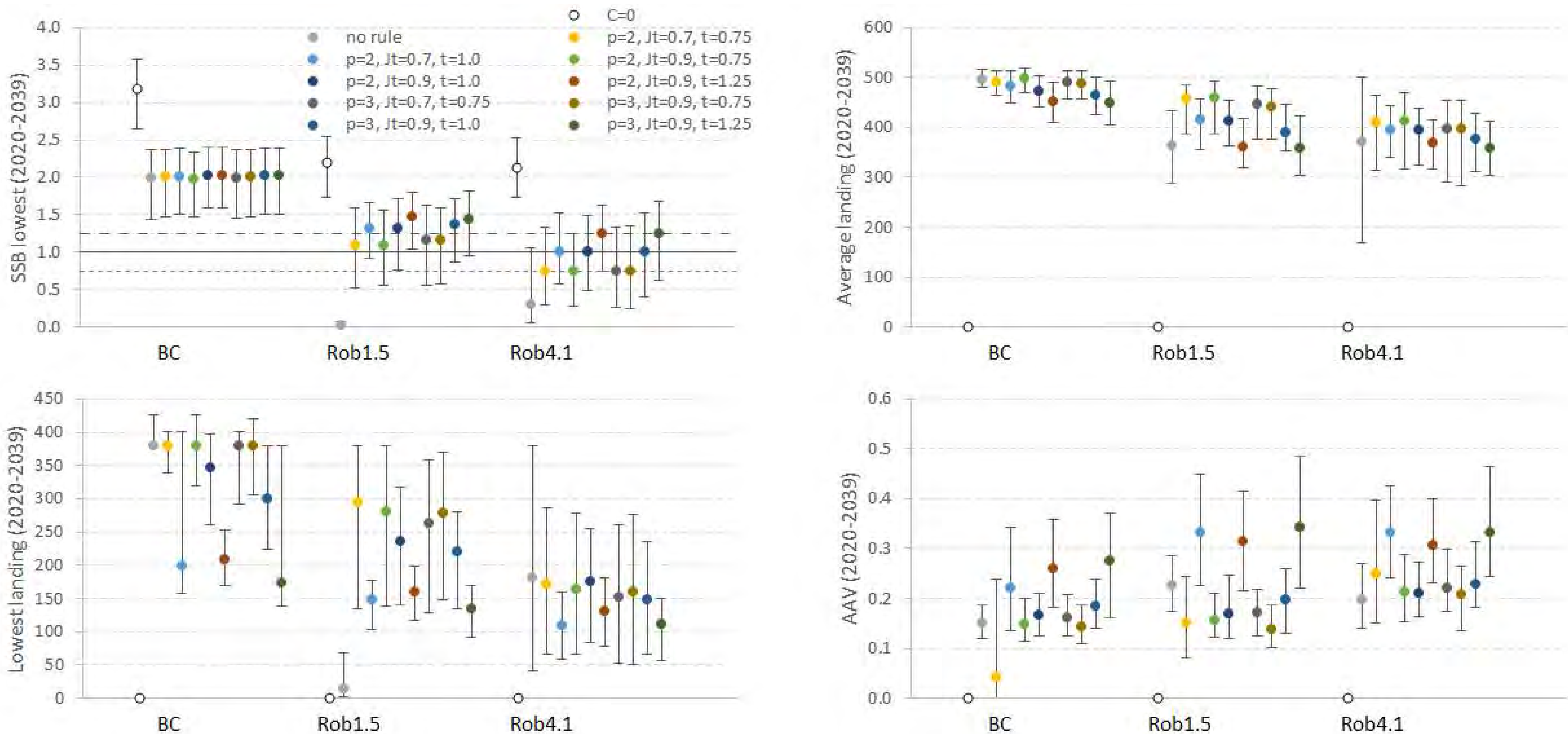


Figure 3: Performance statistics for no landings, no harvest control rule, and MP variants. Gamma values are shown below:

		t=0.75	t=1.0	t=1.25
p=2	Jthresh=0.7	684	293	x
	Jthresh=0.9	666	500	284.5
p=3	Jthresh=0.7	559	x	x
	Jthresh=0.9	580.7	390	222.3

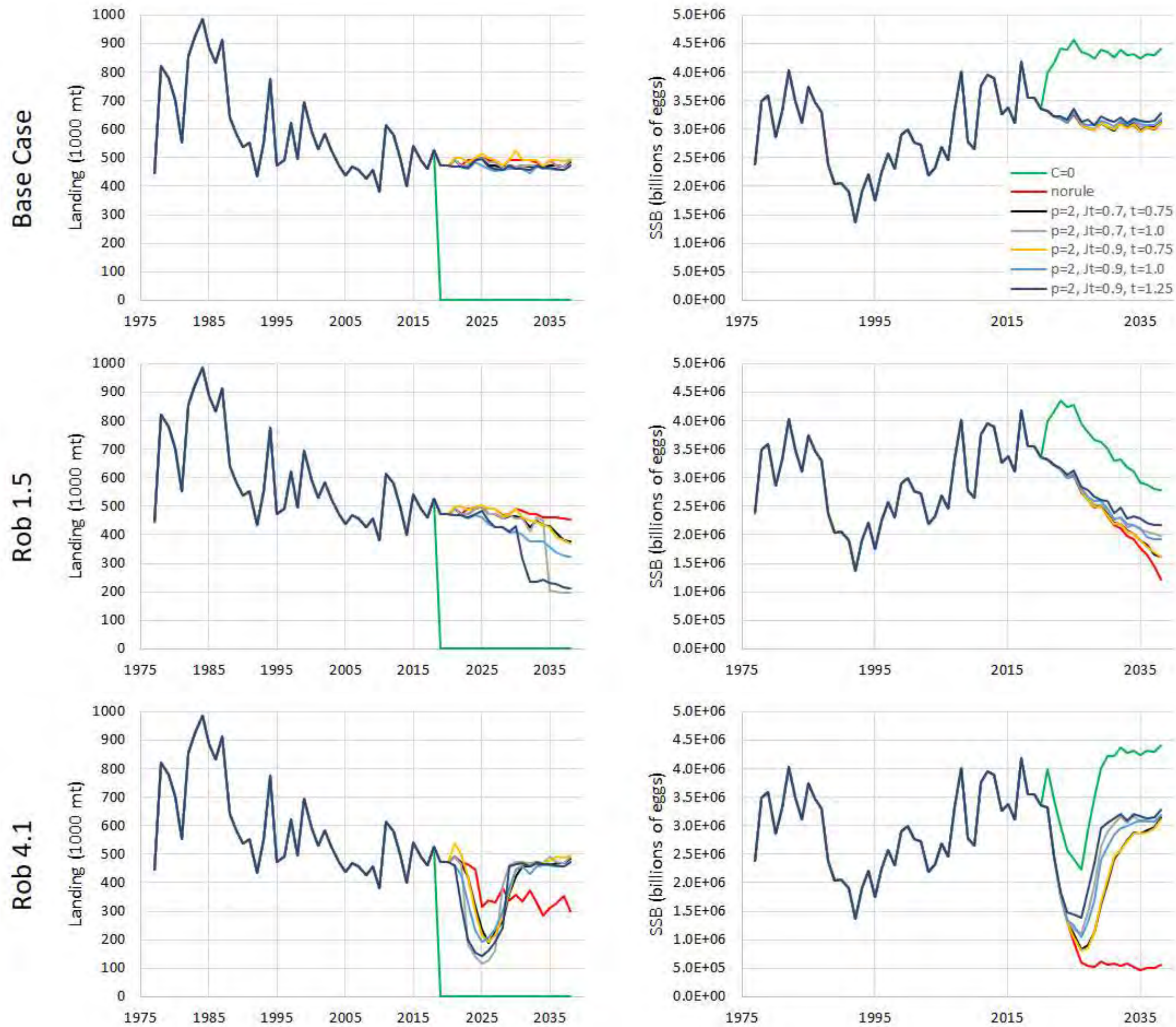


Figure 4a: Historical estimates and projected 20-year for a series of quantities for the **Base Case** OM and **Robustness tests, 1.5 and 4.1**, with no future catch (green lines), without (red lines) and with a series of the MP variants with **p=2**.

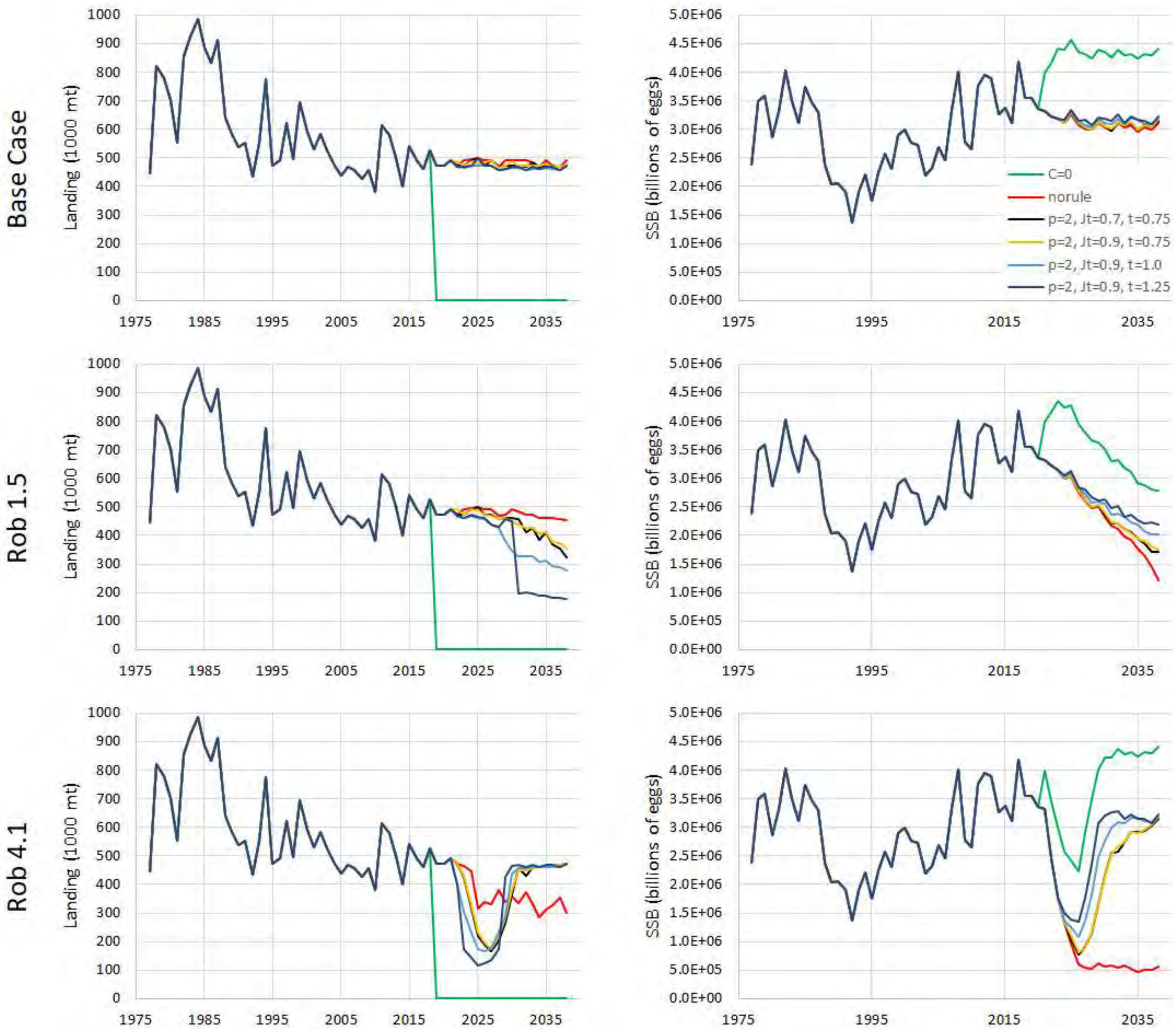


Figure 4b: Historical estimates and projected 20-year for a series of quantities for the **Base Case** OM and **Robustness tests, 1.5 and 4.1**, with no future catch (green lines), without (red lines) and with a series of the MP variants with **p=3.**

Table 1a: Performance statistics for the Base Case OM for a series of MP variants with **p=2**

	no catch			no rule			p=2, Jt=0.7, t=0.75			p=2, Jt=0.7, t=1.0			p=2, Jt=0.9, t=0.75			p=2, Jt=0.9, t=1.0			p=2, Jt=0.9, t=1.25		
	Median	10	90	Median	10	90	Median	10	90	Median	10	90	Median	10	90	Median	10	90	Median	10	90
Related to catch																					
Average landing 2020-2039	0.0	0.0	0.0	494.7	479.2	517.5	491.4	464.2	514.2	481.8	448.6	513.6	499.0	469.0	518.8	472.3	440.8	502.4	451.2	410.6	490.5
Av landing no rule	0.0	0.0	0.0	494.7	479.2	517.5	495.2	479.5	519.5	494.7	479.0	515.9	495.5	473.5	520.4	496.3	476.4	517.4	495.3	476.7	518.3
Av landing with rule	0.0	0.0	0.0	0.0	0.0	0.0	381.0	0.0	461.5	168.0	0.0	198.0	504.8	394.3	564.3	390.0	330.1	426.7	223.7	190.8	248.6
Lowest landing (2020-2039)	0.0	0.0	0.0	379.9	379.9	425.6	379.9	338.6	400.7	198.0	157.0	400.7	379.9	319.8	425.6	347.5	261.6	397.2	208.8	169.0	252.5
2020 landing	0.0	0.0	0.0	473.7	400.7	590.8	473.7	400.7	590.8	473.7	400.7	590.8	473.7	400.7	590.8	473.7	400.7	590.8	473.7	400.7	590.8
Related to abundance																					
Egg(2020)	4.00	3.37	4.79	3.32	2.73	4.10	3.32	2.73	4.10	3.32	2.73	4.10	3.32	2.73	4.10	3.32	2.73	4.10	3.32	2.73	4.10
Egg(2040)	4.50	3.63	5.32	3.17	2.21	4.04	3.18	2.33	4.04	3.21	2.35	4.04	3.19	2.33	4.00	3.20	2.37	4.07	3.22	2.37	4.08
Egg lowest (2020-2039)	3.19	2.66	3.59	2.00	1.44	2.38	2.00	1.48	2.38	2.01	1.51	2.39	1.97	1.47	2.34	2.02	1.60	2.40	2.03	1.60	2.40
Prob Egg(2040) lowest	2			6			6			5			5			5			5		
Related to catch variability																					
AAV 2020-2039	-			0.15	0.12	0.19	0.15	0.12	0.20	0.22	0.14	0.34	0.15	0.11	0.20	0.17	0.12	0.21	0.26	0.18	0.36
AAV with rule	-			-			0.04	0.00	0.24	0.34	0.00	0.63	0.13	0.05	0.26	0.16	0.05	0.24	0.34	0.20	0.50
LAgill																					
Fraction years rule applied	0.00	0.00	0.00	0.00	0.00	0.00	0.05	0.00	0.20	0.05	0.00	0.15	0.25	0.05	0.50	0.20	0.05	0.40	0.15	0.05	0.30
True negative	0.0			0.0			6.5			4.3			24.9			19.4			15.3		
Fals negative	0.0			0.0			1.6			1.4			2.5			2.6			2.3		
False positive	0.0			0.0			3.7			3.4			11.7			11.4			10.7		
True positive	100.0			100.0			88.3			91.0			61.0			66.8			71.8		
Prob rule in 2020	0			0			0			0			0			0			0		
Fraction years Hit Fmax	0	0	0	0	0	0.15	0	0	0.1	0	0	0.1	0	0	0.1	0	0	0.05	0	0	0.05
Hit Fmax, landings not taken	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Table 1b: Performance statistics for the Base Case OM for a series of MP variants with **p=3**

	no catch			no rule			p=3, Jt=0.7, t=0.75			p=3, Jt=0.9, t=0.75			p=3, Jt=0.9, t=0.75			p=3, Jt=0.9, t=1.0		
	Median	10	90	Median	10	90	Median	10	90	Median	10	90	Median	10	90	Median	10	90
Related to catch																		
Average landing 2020-2039	0.0	0.0	0.0	494.7	479.2	517.5	490.3	455.9	513.6	488.0	456.4	513.6	464.3	426.2	500.7	450.4	404.8	492.6
Av landing no rule	0.0	0.0	0.0	494.7	479.2	517.5	495.1	479.0	518.0	495.8	473.2	519.1	494.7	473.9	518.3	494.3	478.2	517.4
Av landing with rule	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	370.6	447.4	0.0	499.2	308.5	0.0	338.5	178.2	0.0	195.3
Lowest landing (2020-2039)	0.0	0.0	0.0	379.9	379.9	425.6	379.9	291.6	400.7	379.9	305.6	421.3	299.5	224.7	379.9	173.3	137.5	379.9
2020 landing	0.0	0.0	0.0	473.7	400.7	590.8	473.7	400.7	590.8	473.7	400.7	590.8	473.7	400.7	590.8	473.7	400.7	590.8
Related to abundance																		
Egg(2020)	4.00	3.37	4.79	3.32	2.73	4.10	3.32	2.73	4.10	3.32	2.73	4.10	3.32	2.73	4.10	3.32	2.73	4.10
Egg(2040)	4.50	3.63	5.32	3.17	2.21	4.04	3.18	2.33	4.04	3.18	2.37	4.04	3.22	2.37	4.05	3.25	2.44	4.09
Egg lowest (2020-2039)	3.19	2.66	3.59	2.00	1.44	2.38	2.00	1.45	2.38	2.00	1.48	2.38	2.02	1.51	2.39	2.02	1.51	2.39
Prob Egg(2040) lowest	2			6			5			5			5			5		
Related to catch variability																		
AAV 2020-2039	-			0.15	0.12	0.19	0.16	0.13	0.21	0.14	0.11	0.19	0.18	0.14	0.24	0.28	0.16	0.37
AAV with rule	-			-			0.00	0.00	0.25	0.11	0.00	0.19	0.20	0.00	0.34	0.34	0.00	0.60
LAgill																		
Fraction years rule applied	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.20	0.20	0.00	0.40	0.15	0.00	0.35	0.15	0.00	0.30
True negative	0.0			0.0			5.4			20.3			16.3			13.8		
Fals negative	0.0			0.0			0.5			1.3			1.2			1.2		
False positive	0.0			0.0			2.9			11.6			10.1			9.4		
True positive	100.0			100.0			91.3			66.9			72.5			75.7		
Prob rule in 2020	0			0			0			0			0			0		
Fraction years Hit Fmax	0	0	0	0	0	0.15	0	0	0.1	0	0	0.1	0	0	0.05	0	0	0.05
Hit Fmax, landings not taken	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Table 2: Median fraction of years the rule is applied for the Base Case and Robustness tests 1.5 and 4.1 for a series of MP variants.

Base Case		t=0.75	t=1.0	t=1.25
p=2	Jthresh=0.7	0.05	0.05	x
	Jthresh=0.9	0.25	0.20	0.15
p=3	Jthresh=0.7	0.00	x	x
	Jthresh=0.9	0.20	0.15	0.15

Rob 1.5		t=0.75	t=1.0	t=1.25
p=2	Jthresh=0.7	0.35	0.25	x
	Jthresh=0.9	0.63	0.60	0.45
p=3	Jthresh=0.7	0.30	x	x
	Jthresh=0.9	0.60	0.50	0.43

Rob 4.1		t=0.75	t=1.0	t=1.25
p=2	Jthresh=0.7	0.40	0.30	x
	Jthresh=0.9	0.60	0.50	0.40
p=3	Jthresh=0.7	0.40	x	x
	Jthresh=0.9	0.55	0.45	0.40

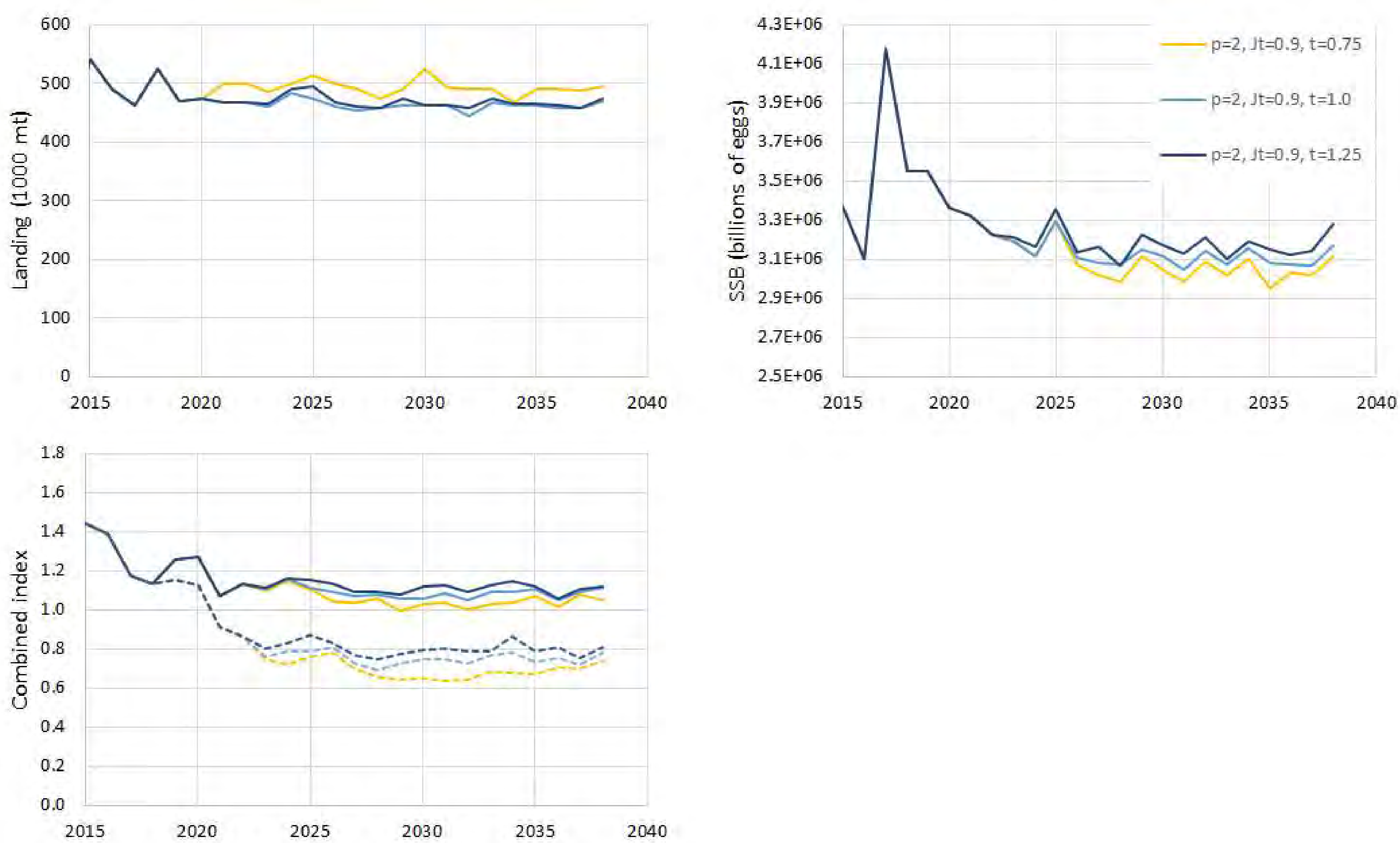


Figure 5: Historical estimates and projected 20-year for a series of quantities for the **Base Case** OM for a series of MP variants with Jthreshold=0.9, p=2 and varying tuning value t**.**

Appendix C - Trade_off_plots_v2

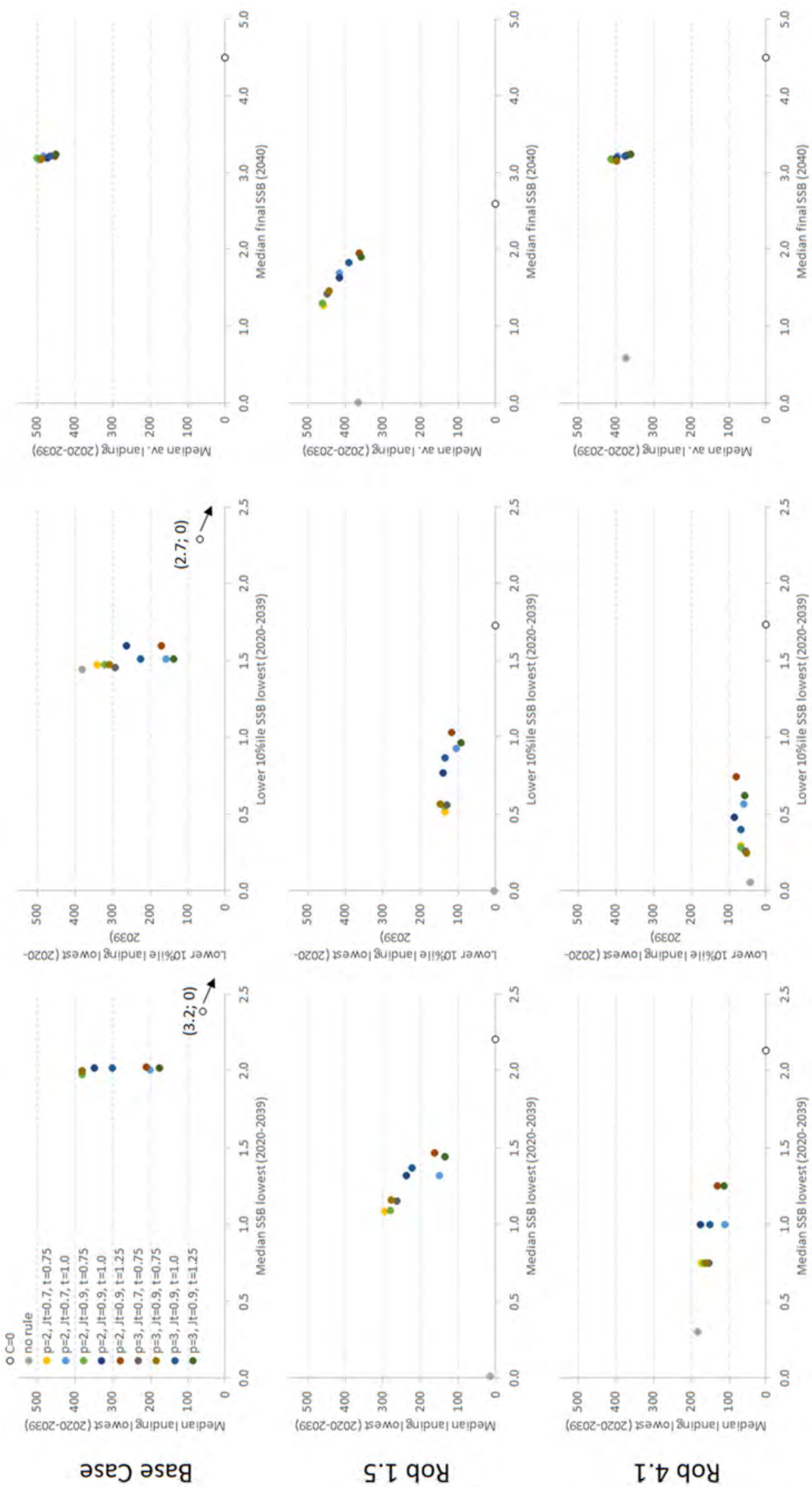


Figure 1: Trade-off plots for the Base Case, and Robustness tests 1.5 and 4.1 under no catch, no rule and a series of MP variants

The 2018 Operational Management Procedure for the South African *Merluccius paradoxus* and *M. capensis* Resources

A. Ross-Gillespie, D.S. Butterworth, J.P Glazer and T.P. Fairweather

[Note that this document is an update of FISHERIES/2014/OCT/SWG-DEM/64 for OMP-2014, and borrows from the text by the authors thereof.]

Introduction

The algorithm for the 2018 Operational Management Procedure (OMP) to provide TAC recommendations for the South African *Merluccius paradoxus* and *M. capensis* resources is empirical. It calculates an increase or decrease of the TAC in relation to the level of an index combining recent CPUE and survey abundance estimates compared to a target level for that index. The basis for the associated computations is set out below, with the tuning parameters given in Table 1.

The 2018 OMP

The species-combined TAC in year *y*+1 is given by:

$$TAC_{y+1} = C_{y+1}^{para} + C_{y+1}^{cap} \tag{1}$$

with

$$C_{y+1}^{spp} = b^{spp}\left(J_y^{spp} - J_0^{spp}\right) \tag{2}$$

where

TAC_y is the total TAC recommended for year *y*,

C_y^{spp} is the intended species-disaggregated TAC for species *spp* year *y*,

J_0^{spp} and b^{spp} are tuning parameters (see Table 1), and

J_y^{spp} is a measure of the immediate past level in the abundance indices for species *spp* that is available to use for calculations for year *y.*

Measure of recent abundance level

The measures of the immediate past level J_y^{spp} for the abundance indices are computed as follows (note that these *J* indices reflect averages over the most recent three years for which the data in question are available):

$$J_y^{para} = \frac{1.0J_y^{WC\_CPUE,para} + 0.75J_y^{SC\_CPUE,para} + 0.5J_y^{WC\_surv,para} + 0.25J_y^{SC\_surv,para}}{2.5} \tag{3}$$

$$J_y^{cap} = \frac{1.0J_y^{WC\_CPUE,cap} + 0.75J_y^{SC\_CPUE,cap} + 0.5J_y^{WC\_surv,cap} + 1.0J_y^{SC\_surv,cap}}{3.25} \tag{4}$$

with

$$J_y^{WC/SC\_CPUE,spp} = \sum_{y'=y-3}^{y-1} I_y^{WC/SC\_CPUE,spp} \Bigg/ \sum_{y=2010}^{2012} I_y^{WC/SC\_CPUE,spp} \quad (5)$$

$$J_y^{WC/SC\_surv,spp} = \sum_{y'=y-2}^{y} I_y^{WC/SC\_surv,spp} \Bigg/ \sum_{y=2011}^{2013} I_y^{WC/SC\_surv,spp} \quad (6)$$

Thus, the weighting of the different indices (denoted by *I*) is taken to be the same as for OMP-2010 and OMP-2014, and the normalization is such that a value of *J*=1 reflects resource abundance about the same as in 2011/2012.

Constraints on TAC change

The maximum allowable annual increase in TAC is 10%, and the maximum allowable annual decrease in TAC is 5% unless the *M. paradoxus* average biomass index falls too low, in which case the maximum allowable annual decrease becomes:

$$MaxDecr_y = \begin{cases} 5\% & if\ J_y^{para} > J^{thresh1,para} \\ \text{linear between } x\% \text{ and } 5\% & if\ J^{thresh2,para} < J_y^{para} < J^{thresh1,para} \\ x\% & J_y^{para} < J^{thresh2,para} \end{cases} \quad (7)$$

x, $J^{thresh1,para}$ and $J^{thresh2,para}$ are tuning parameters (see Table 1).

Further, if J_y^{cap} drops below $J^{thresh,cap}$, then action will be taken to reduce the anticipated catch of *M. capensis* further, probably through measures to have the offshore trawl fishery more in deeper waters as a further TAC drop in these circumstances might reduce the catch of *M. paradoxus* unnecessarily.

Two further constraints are included in OMP-2018:

i. An upper cap on the TAC is imposed, so that the TAC cannot exceed 160 000 tons.
ii. The TACs for 2019 and 2020 are fixed at 146 431 tons.

Procedures in the event of missing data

CPUE data

Non-availability of data to compute the GLM-standardised CPUE series for each species is not anticipated.

Survey data

a) If for one survey at most two years of the most recent three have been missed, the computations continue as indicated, with the missing data omitted from computation of the measures of the immediate past level (equation 6).

b) If all of the most recent three years have been missed (i.e. no data available to compute $J_y^{WC/SC\_surv,spp}$), the level for that index will be ignored in computing the average recent level (equations 3 and 4), but an OMP review will commence immediately.

c) The development of OMP-2018 assumed that the surveys will be conducted by the *Africana* from 2019 onwards, and that for recent pre-2019 surveys conducted by the commercial vessels, those vessels were equivalent to *Africana* in terms of trawling efficiency (catchability coefficient *q*). However, if the *Africana* is unable to conduct some future demersal surveys which provide OMP input, abundance estimates from commercial vessels for those surveys will be multiplied by 1.25 prior to input to equations 3 and 4. (This calibration factor, with its standard error, was estimated from assessments, and the OMP checked for robustness to such a replacement.)

Acknowledgements

Computations throughout the development of OMP-2018 were performed using facilities provided by the University of Cape Town's ICTS High Performance Computing team: hpc.uct.ac.za

Table 1: Tuning parameters for OMP-2018

	M. paradoxus	*M. capensis*
J_0	0.132	0.240
b	88.02	35.00
$J^{thresh1,para}$	0.75	
$J^{thresh2,para}$	0.65	
$J^{thresh,cap}$		0.60
x	25	

Appendix A

Extraction and processing of demersal trawl catch and effort data

A1. Data extraction

Hake catches are reported in two ways:

i) Fine scale data: On the vessel the skipper estimates the catch for each drag, as well as recording important information on depth, longitude and latitude, time and effort [called the "drag" data].
ii) Onshore when the vessel is offloaded (called a landing), catches are more accurately measured for each product category [called the "landing" data]. Each landing is associated with a number of drags made at sea.

When a hake vessel returns from a fishing trip the vessel lands and the catch is discharged to a shore-based processing establishment. The discharged catch for some product categories is graded by size (weight) into product size categories. The catch per product size category is weighed and the total mass (landed_mass) is recorded on the landing sheet. A landing consists of more than one drag (trawl) and the catch estimates per drag are derived from a skipper's estimate made while at sea. At Branch Fisheries the landing is captured first in order to keep track of how much of the TAC has been caught. The captured landing data are then proof-read before the drags are captured. There are 242 species and category codes used in the database of which 59 are for hake alone. A procedure called *Convert to Real Mass* (CRM) is run at the close of each day and when a landing is updated. This procedure scales actual landed mass values to correspond with cleaned mass estimates (for the trip) and then calculates a nominal mass using a raising factor for each species and category code. If a species and category code exists in the landing but not in any of the drags (e.g. skipper only estimates for catch of large hake but factory produces large and medium) then that category is assigned to a table known as drags-no-effort (dne) as it is essentially fish that were landed but not caught.

The input data set used in the CPUE GLM analysis is based on the drag data which are modified in such a way so that the catches (by tonnage) are scaled to reflect the more accurate measures of catch contained in the landing data. The extraction of the drag data (scaled to reflect the landed catches) may result in certain data being excluded, particularly with respect to the data post-2000. Such exclusions arise for the following reasons:

a) Some of the landing records could not be matched perfectly with the associated drag files due to mismatched product codes. If this problem occurred, then all drag records associated with that landing were excluded from the GLM input drag data.
b) Not all category codes were included in the data extracts.
c) The GLM input drag data often in recent years has excluded drags which had no catch associated with them. In large part this reflects the freezer vessels which generally report what is referred to as "daily tallies" where they report all the catch for one day against the last drag of the day. These drag records are flagged as daily tallies in the database to distinguish from drag tally records. As these fishing trips usually last 30 days with at least 3/4 trawls per day the number of drags without catch can be appreciable. How this came to pass is unclear as not all drags without catch were omitted from the previous GLM input drag data when compared with the full database.

In order to improve the percentage of data included in the GLM input the following was done:

- A file containing all the drags that are omitted from the final input to the GLM was created (called non-input drag file)
- A file containing all the landings that could not be matched to drag files was created (called non-input landing file)
- At the non-input landing level, sum hake to get the total hake catch for that landing (Lhake)
- In the non-input drag file, at the drag level, sum hake to get the total hake per drag
- Apportion Lhake across the drags of the non-input drag file in a pro-rata basis to create a new total hake per drag
- Use size structure proportions per season/area/depth to split the total hake catch per drag into small, medium and large hake. These proportions were derived from the data for which items a – c above did not apply, and are simply the proportions of small, medium and large hake within a given cell which, for each year, is defined by a depth range, latitude range (for the West Coast) or longitude range (for the South Coast), and quarter (Jan-Mar, Apr-Jun, July-Sept and Oct-Dec). The reason for defining cells at a quarterly level rather than a monthly level was to avoid getting cells which had no or very few samples in them. Even at the quarterly level there was a need to aggregate across latitude (or longitude) within some depth ranges to ensure sample sizes in each cell greater than or equal to 5.

This process allows for the non-mapped landings to be included in the GLM analyses.

Prior to the application of the procedure to allow for non-mapped landings to be included in the GLM analyses, a number of data exclusions are applied. These are as follows:

1. Exclude all landings where there is only one drag.
2. Exclude all landings where SizedHake = ∑ (HGSml + HGMed + HGLar) = 0
3. Exclude all landings which have fillets in the corresponding dne records
4. Exclude all landings where drag∑HGLar = 0 and dnePQ > 0
5. Exclude all landings where dneSizedHake = 0
   (HakeFillets = FilSml + FilMed + FilUng is calculated but NOT excluded)
6. Exclude all landings where ∑Hake=0
7. Distribute dnePQ into the HGLar column across the drags and add the value to Hake, also dd the HakePQ using the formula HGLar + dnePQ * HGLar/∑HGLar +HakePQ
8. Exclude all drags which have SizedHake = 0 and HGUng>0
9. Distribute HGUng over HG Size (e.g. HGSml + HGSml/SizedHake *HGUng)
10. Distribute dneHGUng and dneBroken over HG Size (e.g. HGSml + HGSml/SizedHake *dneHGUng +dneBroken)
11. Exclude all drag_ID where grid > 899
12. Exclude all drag_ID where effort ≤ 0

There were a number of cases in the drag data where ungraded hake was positive, but the small, medium and large size categories all had zeros recorded. These are erroneous and such drags (and not the entire landing) were deleted.

A2. Data accumulation

Because of the practice of daily tallies the data are accumulated on a daily basis for each vessel before attempting GLM analyses.

The following criteria were adopted for accumulating the database:

- If fishing took place in more than one Division (see Table A1 for explanation of Division) within a day for a particular vessel, the data were allocated to the Division in which at least 2/3 of the drags took place. If a 2/3 majority was not achieved, the records were ignored.

- Different net mesh sizes[1] (75mm, 85mm and 110mm) may have been used on a day. If this occurred, the net mesh size which was used on least 2/3 of the drags for any given vessel was allocated to that day. If there was no two thirds majority, the mesh size was recorded as missing. Two records in the database had a mesh size of zero recorded. In both cases, 110mm was used on all other trawls of the day. Therefore a mesh size of 110mm was assumed for those two records.
- If hake was the recorded target species on at least 2/3 of the drags then the day was recorded as hake-targeted, otherwise it was recorded as non-hake targeted.
- If no depth was recorded for a particular drag (i.e. depth = 0 or 999), it was assumed to be the average depth of the other drags on that day for that particular vessel.
- If fishing took place in two Divisions on one day, the average latitude and longitude pertains only to the latitude and longitude recorded for the dominant Division.
- Namibian and foreign vessels (vessel code $\geq$ 500) were excluded from the accumulated file.

Hence, for a particular vessel, the Demersal database was accumulated over a day, summing over the catches and effort, averaging over depth, latitude and longitude, and including the Division, target species and net mesh size as determined by the decision criteria above.

The analyses are further restricted to offshore companies, a list of which is provided in Table A2.

A3. Identifying potential errors

It is possible that recording errors (typo's) may occur in the DAFF demersal catch database, and an objective means of identifying and excluding erroneous records from the analyses is required. This is achieved by applying a "99% quantile rule". Within the accumulated data, any records (days) where the hake CPUE or by-catch CPUE values exceeded the annual 99% quantile for each CPUE respectively are excluded from the analysis. In addition, any effort values that exceed 1090 minutes on the West Coast and 865 minutes on the South Coast are considered to be potential "mistakes" and are also excluded from the analysis.

A number of records in the accumulated database had positive effort, but zero total catch (i.e. hake + all bycatch species) recorded. It was assumed that these records reflected an aborted drag for some reason or another, and they were therefore excluded from the analyses.

Since the analyses are concerned with the hake stocks, only those days on which hake was recorded as the target species were included in the analyses.

[1] The net mesh size reported in the database refers to the net mesh size that was legally allowed, and not the size that was actually used. New log books that were phased in during 2004 makes allowance for skippers to record the actual mesh size used. Some skippers however continue to record the legal limit for their permit, and not the actual mesh size used. Industry made extensive use of liners in the late 1970s and in the 1980s (and perhaps even in the 1990s), thereby greatly reducing the mesh size. Although Industry recently provided a range of possible years over which the use of liners was believed to have been phased out, the diversity of this range precludes this information from being used in any quantitative manner.

TABLE A1: The drag information extracted from the demersal database to be used in the GLM analysis.

Company code (a code assigned to each fishing company for identification purposes)
Vessel code (a unique code assigned to each fishing vessel for identification purposes)
Power factor (as crudely calculated in the early 1970s)
Vessel class (vessels are assigned to broad categories according to their gross registered tonnage)
Landing date (date on which the catch was landed at port)
Drag date (date on which a drag took place)
Start time (time (hour and minutes) at which drag started)
Effort (the amount of time net was dragged; recorded in minutes)
ICSEAF Division (identifying the Division in which the catch took place – Division 1.6 refers to the West Coast, and Divisions 2.1 and 2.2 refer to the South Coast)
Grid block in which catch was taken (the fishing grounds are divided into 20 minute squares so that catch positions can be reported accurately)
Depth at which catch was taken
Mesh size used (75mm, 85mm or 110mm)
Species targeted[2]
Total hake[3] catch (kg)
Total horse mackerel[3] (*Trachurus capensis*) catch (kg)
Total monk[3] (*Lophius vomerinus*) catch (kg)
Total kingklip[3] (*Genypterus capensis*) catch (kg)
Total East Coast sole[3] (*Austroglossus pectoralis*) catch (kg)
Total West Coast sole[3] (*Austroglossus microlepis*) catch (kg)
Total snoek[3] (*Thyrsites atun*) catch (kg)
Total mackerel[3] (*Scomber japonicus*) catch (kg)
Total white squid[3] (*Loligo vulgaris reynaudii*) catch (kg)
Total red squid[3] (*Todapopsis eblanae*/*Todarodes angolensis*) catch (kg)
Total catch (kg) of other species[4] (e.g. as ribbon fish (*Lepidopus caudatus*) and panga (*Pterogymnus laniarius*))
Amount of hake (kg) which make up the large hake size category
Amount of hake (kg) which makes up the medium hake size category
Amount of hake (kg) which makes up the small hake size category
Amount of hake (kg) which makes up the ungraded hake category
Amount of hake (kg) which makes up the small fillets hake category
Amount of hake (kg) which makes up the medium hake fillets category
Amount of hake (kg) which makes up the ungraded hake fillets category
Amount of hake (kg) which makes up PQ hake category

---

[2] Analyses are restricted to drags/days indicated as hake-directed. However, this field was not completed consistently, so that many indications of "hake direction" in fact reflected effort directed at other species. Although hake is generally the dominant species in the catch and the primary target in most trawls, fishermen often fish in areas or use methods that maximize the catch of certain by-catch species, with a resultant decrease in the hake catch rate. These drags are usually also recorded as hake directed.

[3] Space is provided in the log books for declaring the amount of each of these species caught. Apart from hake, the other species are referred to as declared by-catch.

[4] Space was not provided in the old log books for declaring the catch of these species. The catch of each of these species was determined only at the landing site, and apportioned across the drags of the trip in the same ratio of the catch of targeted species across drags. These species are therefore referred to as undeclared by-catch. The new logbooks (phased in during 2004) provide for the recording all possible species caught per drag.

Latitude position at which catch was taken (minutes have been converted to decimalized minutes)
Longitude position at which catch was taken (minutes have been converted to decimalized minutes)

TABLE A2: The company codes of the offshore companies included in the GLM analyses.

Company Codes			
1	112	144	185
2	113	153	187
3	114	154	188
27	115	155	189
35	117	156	190
36	118	157	191
46	119	158	192
54	120	159	193
55	121	160	194
56	122	161	195
61	123	162	196
62	126	163	197
63	127	164	198
68	128	166	199
69	129	167	200
70	130	168	201
100	131	169	202
101	132	170	203
102	133	171	204
103	134	172	205
104	136	173	206
105	137	174	207
106	138	175	210
107	139	176	211
108	140	178	212
109	141	182	213
110	142	183	
111	143	184	

Appendix B

A summary of the General Linear Modelling approach applied to standardize the CPUE data for the offshore trawl fishery for *Merluccius capensis* and *M. paradoxus* off the coast of South Africa for input to the hake OMP.

B1. Introduction

The models applied to standardize the CPUE data of *Merluccius capensis* and *M. paradoxus* caught offshore off the coast of South Africa are summarised here. This is not straightforward because CPUE indices are required at the species level, but the offshore trawl commercial catch data are routinely recorded only as generic "hake", rather than on a species disaggregated basis. This is because the species are very similar in appearance and can be distinguished only by a trained scientific observer. Consequently algorithms developed by OLRAC (2017), which make use of species proportions by size at depth, as estimated from research surveys and observer records from commercial trips, have been applied to split the hake catches by species at a coast level (west and south) before combining the data from both coasts to perform coast-combined species-specific analyses. Note that this approach can be used from 1978 onwards only, as prior to that the depth of drags was not recorded.

The data used in the analyses are obtained from the demersal database of the Fisheries Branch of the Department of Agriculture, Forestry and Fisheries (DAFF). Appendix A provides a description of the information contained in this database and the process followed to ready the data for analysis purposes.

B2. Separating the species

OLSPS (2017) revised the algorithm utilized in OMP-2014 based upon research and observer data over the period 1985-2017. A GLMM with a logit link function and a binomial distribution was applied. Both west and south coast data were modelled using the equation:

$$P = \frac{1}{1+e^{-\psi}} \qquad \text{(B1)}$$

where:

$$\psi = \mu + \alpha \times depth + \lambda_{position} + \varphi_{sizeclass} + \tau_{sizeclass \times depth} + \theta_{sizeclass \times position} \qquad \text{(B2)}$$

and:

P is the observed proportion of ***M. paradoxus*** by mass for a given trawl,
μ is the model intercept,
$depth$ is the mean depth of the trawl in metres, and α is the associated parameter for the covariate,
$\lambda_{position}$ is a categorical variable, being the latitude bin on the West Coast and the longitude bin on the South Coast,
$\varphi_{sizeclass}$ is a categorical variable for small, medium or large size classes, and
$\tau_{sizeclass \times depth}$ is the interaction between size class and depth
$\theta_{sizeclass \times position}$ is the interaction between size class and position.

This model (Model A6b of Glazer *et al.,* 2018) was selected from a suite of models that differed in terms of input data and explanatory variables (Glazer *et al.*, 2018). The parameter values estimated for this model are provided in Table B1. These will not be updated during the implementation period of the OMP.

The GLMM was run without any record specific weighting. This means that the dependent value for each record is the observed mass proportion of *M. paradoxus*.

B3. The General Linear Models

The following two models (equations B3 and B4) are applied to the *M. capensis* and *M. paradoxus* CPUE data respectively:

$$\ell n(CPUE_{M.cap} + \delta) = \alpha + \beta_{year} + \gamma_{depth} + \eta_{Area} + \kappa_{season} + \lambda_{vessel} + v(CPUE_{snoek}) + \nu'(CPUE_{snoek})^2 + \omega(CPUE_{hmack}) + \omega'(CPUE_{hmack})^2 + interactions + \varepsilon \quad \text{(B3)}$$

$$\ell n(CPUE_{M.par} + \delta) = \alpha + \beta_{year} + \gamma_{depth} + \eta_{Area} + \kappa_{season} + \lambda_{vessel} + v(CPUE_{snoek}) + \nu'(CPUE_{snoek})^2 + \omega(CPUE_{hmack}) + \omega'(CPUE_{hmack})^2 + interactions + \varepsilon \quad \text{(B4)}$$

(Note: to avoid clutter, the subscripts "*capensis*" and "*paradoxus*" for the parameters of equations B3 and B4 have been omitted.)

where:

$CPUE_{M.cap}$ is the catch of *M. capensis* per unit of (hake-directed – the recorded data specifies the target species for each trawl) effort,
$CPUE_{M.par}$ is the catch of *M. paradoxus* per unit of (hake-directed) effort,
α is the intercept,
$year$ is a factor with 40 levels (1978-2017) associated with the year effect,
$depth$ is a factor with 8 levels in both the *M. capensis* and *M. paradoxus* models:

- $d1_{wc}$: 0 - 100m
- $d2_{wc}$: 101 - 200m
- $d3_{wc}$: 201 – 300m
- $d4_{wc}$: 301 – 400m
- $d5_{wc}$: > 400m
- $d6_{sc}$: 0 - 100m
- $d7_{sc}$:101 - 200m
- $d8_{sc}$: > 200m

$area$ is a factor with 6 levels in both the *M. capensis* and *M. paradoxus* models:

- $a1_{wc}$: < 31°00S
- $a2_{wc}$: 31°00S - 33°00S
- $a3_{wc}$: 33°00S - 34°20S
- $a4_{wc}$: > 34°20S
- $a5_{sc}$: < 22°00E
- $a6_{sc}$: ≥ 22°00E,

$seas$ is a factor with 4 levels in both the $M.capensis$ and $M.paradoxus$ models:

- Summer: December - February

Autumn: March - May
Winter: June - August
Spring: September - November,

$vessel$ is a factor associated with each individual vessel in the dataset being analyzed (detailed in Appendix A). Note that for the same vessel, different values of this factor may be estimated for *M. capensis* and *M. paradoxus*.

$CPUE_{snoek}$ and $CPUE_{hmack}$ refer to the CPUE of the bycatch species snoek and horse-mackerel respectively (unlike other major by-catch species, these two species tend **not** to co-occur with hake, so that trawls with proportionally larger catches of these two are reflective of some redirection of fishing effort away from hake, of which account needs to be taken in the GLM),

$interactions$ refer to $year \times depth$, $year \times area$ and $depth \times area$ interactions which allow for spatial density patterns which have changed over time, and ε is the error term, assumed to follow a normal distribution.

δ is a (usually small) constant added to the CPUE of the species being modelled to allow for the occurrence of zero CPUE values - here δ is taken to be 10% of the average nominal CPUE of the species being modelled in the respective datasets, and will change each year as the CPUE database is augmented given new data.

B4. Standardizing the CPUE

The introduction of interactions with year requires that the standardized CPUE (assumed to provide an index of local density) be integrated over area to determine an index of abundance. The boundary separating the west and south Coasts is shown in Figure B1 as being from Cape Agulhas to the tip of the Agulhas Bank so that the whole of the major fishing area of Brown's Bank is included in the west coast. The sizes for depth/latitude (west coast) and depth/longitude (south coast) combinations are shown in Tables B2 and B3.

The formula applied to standardize the CPUE for *M. capensis* and *M. paradoxus* respectively is therefore:

$$CPUE_y = \sum_{strata} \left[e^{\{\alpha + \underline{\beta_{year} + \gamma_{depth}} + \eta_{area} + \text{autumn} + \text{median vessel estimate} + \nu(\text{snoek } \overline{CPUE}) + \nu'(\text{snoek } \overline{CPUE^2}) + \varpi(\text{hmack } \overline{CPUE}) + \varpi'(\text{hmack } CPUE^2) + \text{interactions}\}} - \delta \right] * \frac{A_{stratum}}{A_{total}} \quad \text{(B6)}$$

$A_{stratum}$ is the size of the area of the stratum in nm^2 (e.g. depth 200-300m and latitude 31 - 33°), andA_{total} is the total size of the area considered (it is not strictly necessary to divide by A_{total}, but this keeps the units and size of the standardised CPUE index comparable with those of the basic CPUE data).

For the west coast the standardised CPUE is calculated for depths > 200m since very little fishing takes place at depths below 200m. The majority of hauls within the 0 - 200m depth range occur very close to the 200m depth contour, and accordingly are of questionable representativeness of densities within the whole depth-latitude stratum to which the above equation would take them to refer. Similarly, the standardized CPUE for the south coast is calculated for depths > 100m only.

Reference

Glazer JP, Bergh MO, Butterworth DS, Durholtz D and A. Ross-Gillespie. 2018. Further hake specie-splitting algorithm results. Unpublished DAFF Working Group Document: FISHERIES/2018/JULY/SWG-DEM/27. 9pp.

Table B1: Coast-specific parameter values for substitution into equations (B1) and (B2).

	WEST COAST		
	Fish Size		
Parameter	**Small**	**Medium**	**Large**
Intercept	-3.8467	-6.7956	-6.9164
Mean depth	0.0175	0.0206	0.0189
lat<3000	0.0000	0.0000	0.0000
3000≤lat<3100	0.4625	0.6913	0.3755
3100≤lat<3200	0.3872	1.0850	0.9694
3200≤lat<3300	0.3685	1.1997	1.0839
3300≤lat<3400	0.2336	1.2579	1.1591
3400≤lat<3500	-0.0583	0.9015	0.9161
lat≥3500	0.2276	1.0643	0.7884

	SOUTH COAST		
	Fish Size		
Parameter	**Small**	**Medium**	**Large**
Intercept	-6.2389	-6.4230	-6.5111
Mean depth	0.0176	0.0167	0.0149
long<2100	0.0000	0.0000	0.0000
2100≤long<2200	1.4186	1.3607	1.0231
2200≤long<2300	1.2845	1.1622	1.0650
2300≤long<2400	2.5010	2.0649	1.5846
2400≤long<2500	2.4171	2.2306	1.8074
2500≤long<2600	2.1252	1.6150	1.4271
long≥2600	1.6083	0.6589	0.8975

Table B2: The sizes of the areas (nm^2) covered by each of the latitude/depth combination strata on the West Coast.

	Depth (m)		
Latitude (S)	201-300	301-400	401-500
≤31^{o}00	3598 (10.3)	801 (2.3)	657 (1.9)
31^{o}00-33^{o}00	2842 (8.1)	2383 (6.8)	1427 (4.1)
33^{o}00-34^{o}20	882 (2.5)	458 (1.3)	501 (1.4)
>34^{o}20	1357 (3.9)	726 (2.1)	586 (1.7)

Table B3: The size of the area (nm^2) covered by longitude/depth combinations on the South Coast.

	Depth (m)	
Longitude (E)	101-200	201-500
< 22^o	6911 (19.8)	839 (2.4)
≥22^o	8470 (24.2)	2535 (7.2)

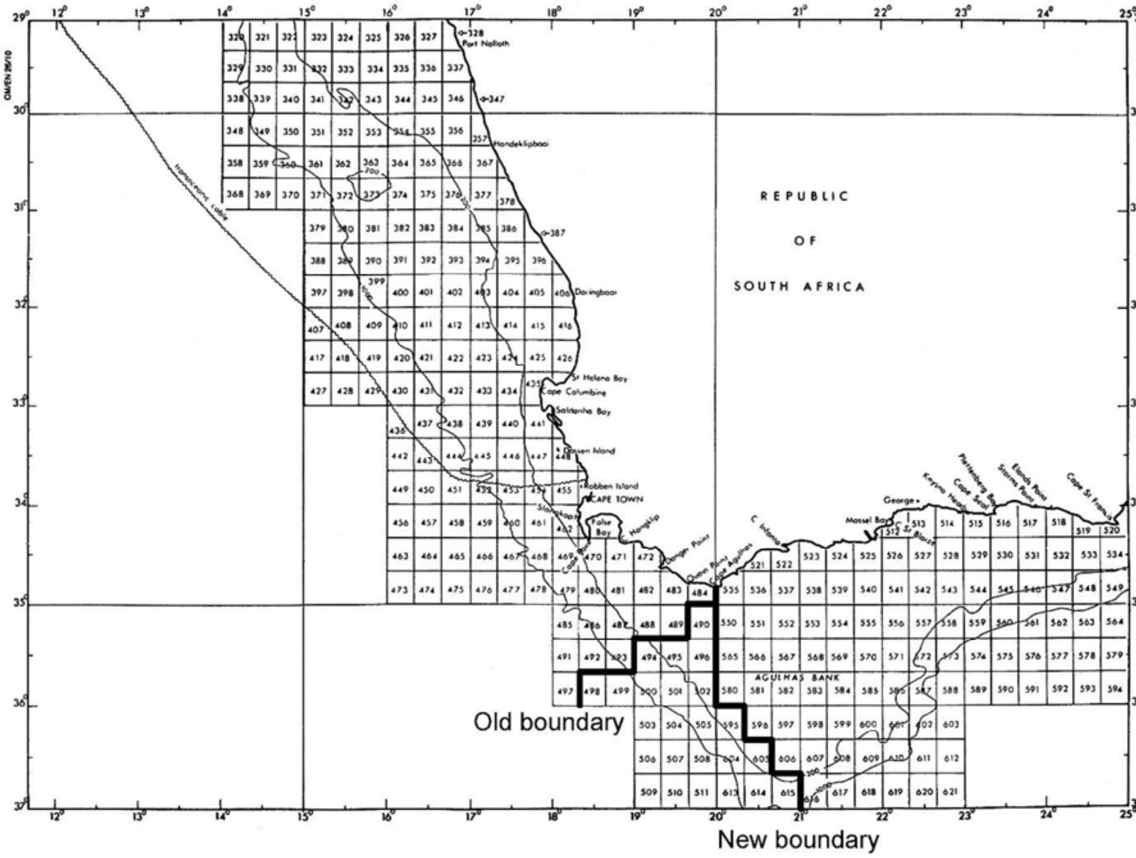


Figure B1: Demarcation of boundaries separating the west and south coasts in the hake fishery. The "Old boundary" was set by ICSEAF and was used to separate coasts until 2004 after which it was agreed by the Demersal Working Group to adopt the "New boundary" for future analyses so that the boundary did not split Brown's Bank. The depth contours shown are the 200m and 1000m contours respectively.

Appendix C

Demersal Research Surveys – sampling strategy, data collection, raised length frequencies and calculation of abundance estimates as applied to Cape hakes (*Merluccius capensis* & *M. paradoxus*)

C1. Survey Design

Demersal surveys cover the same geographical range each year. West Coast surveys extend from the coast out to the 500 metre isobath and from the international border between South Africa and Namibia to Cape Agulhas (20° E longitude), while South Coast surveys cover the same depth range from Cape Agulhas to 27° E longitude. Stations are selected using a pseudo-random stratified sampling design. The area is divided into depth strata and each stratum is further subdivided into 1° latitude substrata on the West Coast (Table C1a) and 1° longitude substrata on the South Coast (Table C1b). Stations within each substratum are selected at random, and the number of target stations per substratum is proportional to the area of the substratum.

Table C1a: Area (nm^2) of depth and latitude strata used on the West Coast of South Africa for Demersal Surveys

<table>
<tr><th>Lat\Depth</th><th>000-100</th><th>101-200</th><th>201-300</th><th>301-400</th><th>401-500</th></tr>
<tr><td>28°30-29</td><td>239.27</td><td>312.53</td><td>0</td><td>0</td><td>0</td></tr>
<tr><td>29-30</td><td>345.3</td><td>4098.38</td><td>447.49</td><td>173.26</td><td>252.3</td></tr>
<tr><td>30-31</td><td rowspan="2">687.55</td><td>2301.22</td><td>3150.3</td><td>627.42</td><td>404.82</td></tr>
<tr><td>31-32</td><td>2080.96</td><td>1535.9</td><td>1121.03</td><td>1016.07</td></tr>
<tr><td>32-33</td><td>814.69</td><td>1302.36</td><td>1306.45</td><td rowspan="2">1585.85</td><td rowspan="2">824.19</td></tr>
<tr><td>33-34</td><td>678.16</td><td>860.71</td><td>550.25</td></tr>
<tr><td>34-35</td><td>1244.8</td><td>1366.69</td><td>641.22</td><td rowspan="2">709.32</td><td rowspan="2">521.71</td></tr>
<tr><td>35-36°20</td><td>62.41</td><td>1820.77</td><td>896.65</td></tr>
<tr><td>TOTAL</td><td>4072.18</td><td>14143.62</td><td>8528.26</td><td>4216.88</td><td>3019.09</td></tr>
</table>

Table C1b: Area (nm^2) of depth and longitude strata used on the South Coast of South Africa for Demersal Surveys

Long\Depth	000-050	051-100	101-200	201-500
20-21	303.57	1804.2	3750.72	454.22
21-22	138.06	1930.39	3804.62	839.05
22-23	230.39	2080.29	3389.52	1206.37
23-24	100.36	651.68	1783.61	533.91
24-25	183.39	231.76	1419.01	347.78
25-26	330.65	385.01	978.24	281.79
26-27	206.79	512.61	899.12	164.97
TOTAL	**1493.21**	**7595.94**	**16024.84**	**3828.09**

C2. Gear Type

Surveys conducted on the research vessel *Africana* between 1985 and September 2003 used a 2-panel German 180 ft trawl net with a rope-wrapped chain footrope, 150kg lift and 1500kg WV doors. In 2003, "new" gear was introduced that consisted of a 4-panel German 180 ft trawl net with a modified rockhopper footrope, 150kg lift and 1500kg Morgere multi-purpose doors. The "new" gear has subsequently been used as standard on *Africana* (with the exception of 2006 and 2010, where the old gear was used to facilitate a gear "cross-calibration") and on the fishing vessels *Andromeda* and *Compass Challenger*.

C3. Summary of Demersal Abundance Surveys

West Coast surveys were completed bi-annually (summer and winter) from 1983 to 1990, and in summer only from 1991 onwards (Table C2). The data from the first survey (summer 1983) are not used as this is regarded as a learning or "shake-down" survey. Extensive use was made of bobbin-gear during the 1983 and 1984 surveys, as many of the stations were in areas that were previously un-trawled. From 1985 onwards, bobbin-gear was no longer used (Payne *et al*. 1986). Consequently the abundance estimates from the first two years may not be compatible with the rest of the time-series, as the selectivity of the bobbin-gear differs from that of the footrope-trawl gear used from 1985 onwards. During the summer survey of 1989, the vessel broke down after only 25 stations were completed and the survey was aborted. All surveys subsequent to this were successfully completed with the exception of 1993 (where portions of the inshore strata were not adequately surveyed) and 1998 (during which year no surveys were completed as the *Africana* was undergoing a complete re-fit). In 2000 and 2001 the Norwegian research vessel *Dr Fridjtof Nansen* was used to conduct the surveys but these data are not currently used in hake assessments or OMPs.

The first of the South Coast surveys was completed in spring (September) 1986 and the first autumn (April/May) survey was completed in 1988 (Table C2). The following two autumn surveys were only completed within the 200m depth contour, as were the spring surveys from 1990 to 1995. With the exception of 2001 and 2002, surveys of the entire South Coast shelf up to 500m have been completed every autumn since 1999 (although the *Dr Fridjtof Nansen* was used in 2000). In 2002 the *Africana* resumed operations, completing all surveys until April 2012. The commercial fishing vessel *Andromeda*, was used in 2013 (summer), 2014 (summer and autumn) and 2015 (summer and autumn). The *Andromeda* was unavailable in 2016 and was replaced by the *Compass Challenger* for the summer and autumn surveys. The *Africana* was operational for spring 2016 and summer 2017 before undergoing further major repairs. No demersal surveys were completed in 2018.

Table C2: Summary of abundance estimate surveys completed since 1985. Surveys AFR069, AFR109 and AFR281 were inadequately sampled and several South Coast surveys were completed within the 200m depth contour as opposed to the entire 500m area. Surveys completed on the *Dr Fridjof Nansen* are underlined, *Africana* surveys using "new" gear are in bold and *Andromeda* surveys are both bold and underlined.

	WEST COAST		**SOUTH COAST**	
year	**Summer (Jan)**	**Winter (July)**	**Autumn (April/May)**	**Spring (Sept)**
1985	AFR 028	AFR 033		
1986	AFR 039	AFR 046		AFR 048
1987	AFR 050	AFR 054		AFR 056
1988	AFR 059	AFR 066	AFR 063	
1989	AFR 069	AFR 075	AFR 072 <200m	
1990	AFR 079	AFR 084	AFR 082 <200m	AFR 086 <200m
1991	AFR 088		AFR 093	AFR 095 <200m
1992	AFR 100		AFR 102	AFR 106 <200m
1993	AFR 109		AFR 111	AFR 116 <200m
1994	AFR 118		AFR 122	AFR 125 <200m
1995	AFR 127		AFR 129	AFR 131 <200m
1996	AFR 133		AFR 135	
1997	AFR 139		AFR 144	
1998	NO SURVEYS COMPLETED AS *AFRICANA* NOT OPERATIONAL			
1999	AFR 150		AFR 152	
2000	NAN 001		NAN 003	
2001	NAN 004			AFR 160
2002	AFR 165			
2003	AFR 173		AFR 177	**AFR 182**
2004	**AFR 188**		**AFR 191**	**AFR 200a**
2005	**AFR 203**		**AFR 206**	
2006	AFR 214		AFR 217	AFR 224
2007	**AFR 228**		**AFR 232**	**AFR 236**
2008	**AFR 238**		**AFR 241**	**AFR 246**
2009	**AFR 249**		**AFR 252**	
2010	AFR259		AFR261	
2011	**AFR270**		**AFR273**	
2012	**AFR279**		**AFR281**	
2013	**AND001**			
2014	**AND002**		**AND003**	
2015	**AND004**		**AND005**	
2016	**CCH008**		**CCH009**	
2017	**AFR291**			
2018	NO SURVEYS COMPLETED AS *AFRICANA* NOT OPERATIONAL			

C4. Data collection

Once the trawl is hauled and emptied onto the deck the catch is sorted depending on species and size composition:

1. Catch of mainly demersal species: sort into species to weigh, if necessary the hake (and occasionally other species) are separated into size categories when the catch is bimodal. This is done because the reality of sorting fish is that people are inclined to pick up the bigger fish first and thus the first few bins, if not sorted, would be mainly large fish whereas the last would be mainly small fish and neither will be suitable for a length frequency measurement. In addition, either a sub-sample of or all the hake is sexed, within each size category and the sexed hake are also measured.
2. Catch of mainly pelagic species – mixed sizes: occasionally the trawl will encounter a school of pelagic fish – usually redeye, anchovy or horse mackerel. If the catch is large (>1500kg) and includes a varied size range of demersal species then the demersal species are picked out and separated as discussed above and the pelagic species are weighed and dumped with a sub-sample measure. If the catch is exceptionally large (>2 500kg) then the whole catch will be sub-sampled with half or the majority being dumped as "mix" and a reasonable number of bins sorted and used to scale up the catch amount.
3. Catch of mainly pelagic species – small sizes: catches of small pelagic and demersal fish, usually made in shallower water, are sub-sampled (usually one or two bins) and the ratio is used to scale up to the weight of the dumped mix.

Once sorted to species (and gender and size category where necessary), the total weight of each species (and category where relevant) is recorded. Length frequency data are then recorded for each species (and category) where feasible (in some cases, a count of the number of individuals in the sample is recorded, rather than length measurements). Sub-samples of the "commercial" species, namely hake, monk, kingklip, squid and sole are dissected to determine individual length, weight, sex, maturity, stomach contents and otoliths (or illicia or statoliths) are removed for age determination purposes.

C5. Survey abundance indices

Catch data collected during the surveys is used to calculate an abundance estimate by the swept-area survey method. Two basic assumptions of the swept area method are that all fish in the path of the net are caught, and that the fish are distributed homogeneously over the survey area. Both of these assumptions are open to criticism and are difficult to defend. However, it is reasonable to assume that the effects of these two assumptions will not vary much from year to year. Therefore abundance estimates obtained using the swept area method are not regarded as absolute estimates, but rather as relative abundance indices. The assumption is that each trawl (*j*) within a stratum (*i*) gives an independent estimate of the density in that stratum. Then the average density for all trawls in a stratum will be an estimate of the average density in the stratum. Therefore multiplying the average density (kg/nm^2) by the area of the stratum (nm^2) gives an estimate of the total abundance in that stratum.

1. Calculate the area swept (nm^2) a_{ij} for each trawl: where s_{ij} is the towing speed (knots, nm/hr), t_{ij} is the duration (minutes) and w_{ij} is the horizontal mouth width (m) i.e. the width of the trawl track in the *j*-th trawl of the *i*-th stratum;

$$a_{ij} = s_{ij} \times \frac{t_{ij}}{60} \times \frac{w_{ij}}{1852}$$

2. Calculate the observed density (kg/nm^2) d_{ij} in the *j*-th trawl of the *i*-th stratum for each trawl where C_{ij} is the observed catch weight (kg) of the species and a_{ij} is the area swept (nm^2);

$$d_{ij} = \frac{C_{ij}}{a_{ij}}$$

3. Calculate the mean density (kgs/nm^2) $\bar{d}_i$ per stratum and its standard error $SE(\bar{d}_i)$ where n_i is the number of trawls in the *i*-th stratum and d_{ij} is the observed density in the *j*-th trawl of the *i*-th stratum;

$$\bar{d}_i = \frac{\sum_{j=1}^{n_{ij}} d_{ij}}{n_i} \; ; SE(\bar{d}_i) = \frac{1}{\sqrt{n_i}} \sqrt{\frac{n_i \sum_{j=1}^{n_i} d_{ij}^2 - \left(\sum_{j=1}^{n_i} d_{ij}\right)^2}{n_{i(n_i-1)}}}$$

4. Estimate abundance per stratum B_i (tons) where $\bar{d}_i$ is the mean density (kg/nm^2) and A_i is the area (nm^2) of the i-th stratum, division by 1000 is to get from kg to tons;

$$B_i = \frac{\bar{d}_i \times A_i}{1000}$$

5. The total abundance estimate (tons) for the survey area B is the sum of the abundance per stratum B_i over all strata n_s ;

$$B = \sum_{i=1}^{n_s} B_i$$

6. Multiply the standard error of the mean density per stratum by the area of the stratum area to get estimated standard error per stratum;

$$SE(B_i) = \left(SE(\bar{d}_i) \times A_i\right)$$

7. Sum the square of the standard error per stratum over all strata to get the standard error of the total abundance estimate for the survey area.

$$SE(B) = \sqrt{\sum_{i=1}^{n_s} SE(B_i)^2}$$

where

B is the abundance estimate for the total survey area, $SE(B_i)$ is the standard error of the abundance for the *i*-th stratum and $SE(B)$ is the standard error of the overall abundance estimate.

Survey abundance indices and standard errors are presented in Table C3 for *M. paradoxus* and Table C4 for *M. capensis* – note for both tables the values in bold represent surveys when *Africana* used new gear; surveys conducted on the *Andromeda* and *Compass Challenger* are underlined and bold values and shaded surveys either only extended to 200m or were incomplete and have therefore been omitted.

C6. References

Payne, A.I.L., C.J. Augustyn and R.W. Leslie (1986): Results of the South African hake biomass cruises in Division 1.6 in 1985. *Colln scient. Pap. int. Commn SE. Atl. Fish.* 13(2): 181-196.

Table C3: Survey abundance estimates and associated standard errors (in thousand tons) for *Merluccius paradoxus*. *Africana* surveys using "new" gear are in bold, *Andromeda* and *Compass Challenger* surveys are both bold and underlined and surveys marked in grey were inadequately sampled.

Year	WEST COAST				SOUTH COAST			
	Summer (Jan)		Winter (July)		Autumn (April/May)		Spring (Sept)	
	Abundance	SE	Abundance	SE	Abundance	SE	Abundance	SE
1985	168.989	37.765	290.281	63.295				
1986	202.334	37.745	147.378	21.667			11.280	3.111
1987	284.434	54.165	180.158	39.047			16.381	3.033
1988	138.534	20.303	252.121	71.246	28.293	8.673		
1989			434.092	142.716				
1990	307.615	87.841	205.704	43.607				
1991	331.177	81.633			27.570	8.153		
1992	225.755	33.711			25.036	6.650		
1993	340.079	51.427			162.375	81.691		
1994	333.499	56.259			108.179	38.369		
1995	317.104	76.709			70.890	39.330		
1996	474.270	92.744			68.859	19.929		
1997	543.615	96.043			121.707	51.507		
1998								
1999	542.830	110.541			263.256	59.439		
2000								
2001							16.668	7.159
2002	251.820	32.690						
2003	386.321	63.565			185.345	82.188	**98.434**	**42.249**
2004	**271.540**	**55.710**			**39.822**	**22.153**	**70.001**	**22.156**
2005	**296.065**	**42.409**			**26.691**	**6.017**		
2006	316.247	57.332			34.868	5.843	68.507	18.283
2007	**407.377**	**77.222**			**102.195**	**53.688**	**66.267**	**21.966**
2008	**238.143**	**37.018**			**33.034**	**9.340**	**25.661**	**8.324**
2009	**310.760**	**27.768**			**45.030**	**15.551**		
2010	576.848	88.202			46.938	12.160		
2011	**380.185**	**128.013**			**21.054**	**6.531**		
2012	**405.865**	**59.099**						
2013	**136.260**	**25.116**						
2014	**269.482**	**37.492**			**62.925**	**24.802**		
2015	**207.583**	**24.057**			**111.411**	**51.852**		
2016	**312.876**	**33.250**			**94.177**	**51.731**	**22.520**[5]	**6.700**
2017	**319.024**	**58.766**						
2018								

[5] Note that this survey estimate was inadvertently omitted from the updated assessments on which OMP-2018 is based, but this omission would have had little impact on the results

Table C4: Survey abundance estimates and associated standard errors (in thousand tons) for *Merluccius capensis*. *Africana* surveys using "new" gear are in bold, *Andromeda* and *Compass Challenger* surveys are both bold and underlined and surveys marked in grey were inadequately sampled.

Year	WEST COAST				SOUTH COAST			
	Summer (Jan)		Winter (July)		Autumn (April/May)		Spring (Sept)	
	Abundance	SE	Abundance	SE	Abundance	SE	Abundance	SE
1985	102.929	18.888	159.198	18.982				
1986	113.154	23.474	115.218	19.733			96.768	10.737
1987	75.438	9.709	83.050	10.306			137.008	13.057
1988	66.365	9.930	48.046	9.574	154.548	23.984		
1989			294.740	67.495				
1990	400.142	97.102	156.337	22.507				
1991	67.565	9.656			276.607	25.274		
1992	95.401	11.892			124.495	13.600		
1993	93.613	14.390			144.551	12.379		
1994	124.497	37.845			153.790	20.310		
1995	193.292	24.270			222.464	31.245		
1996	87.969	9.866			222.176	23.144		
1997	252.606	42.721			163.163	17.274		
1998								
1999	188.624	31.362			171.946	13.330		
2000								
2001							117.590	20.093
2002	105.093	16.130						
2003	73.020	12.518			117.538	17.192	**73.604**	**9.142**
2004	**194.294**	**30.714**			**92.796**	**11.318**	**96.933**	**13.936**
2005	**63.363**	**11.498**			**68.672**	**5.302**		
2006	73.655	17.255			116.298	11.931	92.831	8.998
2007	**73.230**	**9.306**			**65.935**	**5.303**	**67.937**	**6.553**
2008	**52.577**	**7.069**			**102.169**	**9.681**	**87.836**	**9.723**
2009	**140.437**	**26.486**			**111.191**	**10.832**		
2010	162.402	34.891			170.261	33.235		
2011	**89.095**	**23.574**			**105.424**	**10.688**		
2012	**84.746**	**8.331**						
2013	**30.383**	**4.575**						
2014	**219.756**	**60.342**			**63.389**	**6.415**		
2015	**65.086**	**9.178**			**76.059**	**6.873**		
2016	**115.058**	**30.400**			**83.197**	**6.600**	**110.301**[6]	**13.436**
2017	**69.289**	**14.486**						
2018								

[6] Note that this survey estimate was inadvertently omitted from the updated assessments on which OMP-2018 is based, but this omission would have had little impact on the results

Appendix D

Procedures for deviating from OMP output for the recommendation for a TAC, and for initiating an OMP review

D1. Metarule Process

Metarules can be thought of as "rules" which pre-specify what should happen in unlikely, exceptional circumstances when application of the TAC generated by the OMP is considered to be highly risky or inappropriate. Metarules are not a mechanism for making small adjustments, or 'tinkering' with the TAC from the OMP. It is difficult to provide firm definitions of, and to be sure of including all possible, exceptional circumstances. Instead, a process for determining whether exceptional circumstances exist is described below (see Fig. D1). The need for invoking a metarule should be evaluated by the DAFF BRANCH FISHERIES [Demersal] Scientific Working Group (hereafter indicated by WG), but only provided that appropriate supporting information is presented so that it can be reviewed at a WG meeting.

D1.1 Description of Process to Determine Whether Exceptional Circumstances Exist

While the broad circumstances that may invoke the metarule process can be identified, it is not always possible to pre-specify the data that may trigger a metarule. If a WG Member or Observer, or DAFF BRANCH FISHERIES Management, is to propose an exceptional circumstances review, then such person(s) must outline in writing the reasons why they consider that exceptional circumstances exist, and must either indicate where the data or analyses are to be found supporting the review, or must supply those data or analyses in advance of the WG meeting at which their proposal is to be considered.

Every year the WG will:

- Review population and fishery indicators, and any other relevant data or information on the population, fishery and ecosystem, and conduct a simple routine updated assessment (likely no more than the core Reference Case model used in the OMP testing refitted taking a further year's data into account).
- On the basis of this, determine whether there is evidence for exceptional circumstances.

Examples of what might constitute an exceptional circumstance in the case of [hake] include, but are not necessarily limited to:

- Survey estimates of abundance that are appreciably outside the bounds predicted in the OMP testing.
- CPUE trends that are appreciably outside the bounds predicted in the OMP testing.
- Catch species composition in major components of the fishery that differ markedly from previous patterns (and so may reflect appreciable changes in selectivity).

Every two years the WG will:

- Conduct an in depth stock assessment (more intensive than the annual process above, and in particular including the full Reference Set of assessment models and conducting of a range of sensitivity tests).
- On the basis of the assessment, indicators and any other relevant information, determine whether there is evidence for exceptional circumstances.

The primary focus for concluding that exceptional circumstances exist is if the population assessment/indicator review process provides results appreciably outside the range of simulated population and/or other indicator trajectories considered in OMP evaluations. This includes the core (Reference case or set of) operating models used for these evaluations, and likely also (though subject to discussion) the operating models for the robustness tests for which the OMP was considered to have shown adequate performance. Similarly, if the review process noted regulatory changes likely to affect appreciable modifications to outcomes predicted in terms of the assumptions used for projections in the OMP evaluations (e.g. as a result, perhaps, of size limit changes or closure of areas), or changes to the nature of the data collected for input to the OMP beyond those for which allowance may have been made in those evaluations, this would constitute grounds for concluding that exceptional circumstances exist in the context of continued application of the current OMP.

(Every year) IF the WG concludes that there is no or insufficient evidence for exceptional circumstances, the WG will:

- Report to the Chief Director Research, DAFF BRANCH FISHERIES that exceptional circumstances do not exist.

IF the WG has agreed that exceptional circumstances exist, the WG will:

- Determine the severity of the exceptional circumstances.
- Follow the "Process for Action" described below.

D1.2 Specific issues that will be considered annually (regarding Underlying Assumptions of the Operating Models (OMs) for the OMP Testing Process)

The following critical aspects of assumptions underlying the OMs for [hake] need to be monitored after OMP implementation. Any appreciable deviation from these underlying assumptions may constitute an exceptional circumstance (i.e. potential metarule invocation) and will require a review, and possible revision, of the OMP:

- Whether selectivities-at-length for the major fisheries differ substantially from assumptions made to generate operating model projections.
- Whether standardised CPUE and survey abundance estimates are within the bounds indicated in operating model projections, where bounds here and in similar cases following shall be taken to be the 5%ile and 95%ile of projections under the Reference Set (RS) of operating models.
- Whether the proportions of *M. capensis* in the west and south coast offshore trawl catches are within the bounds indicated in operating model projections.
- Whether future recruitment levels are within the bounds projected by the RS operating models.
- Whether updates of major data sets or ageing practices indicate substantial differences from what were used to condition the operating models for the OMP testing.
- Whether there have been a series of substantial differences between TACs allocated and the catches subsequently made.
- Whether fishing regulations and/or strategies have changed substantially, and in a manner such that continuing use of the agreed GLM-standardisation procedures would likely introduce substantial bias in resource abundance trend estimates based on CPUE indices.
- Whether new data or information suggest a substantial revision of estimates of stock status or of the spawning biomass at MSY for *M. paradoxus*; the target objective for the fishery is to keep this stock somewhat above its MSY level so that a relatively high CPUE value is maintained (this last for reasons of economic viability).
- Whether updated assessments suggest that the spawning biomass for the *M. paradoxus* population has fallen below its median 2007 level, which will be considered a limit reference point for the fishery.

A guide as to what constitutes "substantial" is a change that would alter the recommended TAC by more than 3%.

D1.3 Description of Process for Action

If making a determination that there is evidence of exceptional circumstances, the WG will with due promptness:

- Consider the severity of the exceptional circumstances (for example, how severely "out of bounds" are the recent CPUEs and survey abundance estimates or recruitment estimates).
- Follow the principles for action (see examples below).
- Formulate advice on the action required (this could include an immediate change in TAC, a review of the OMP, the relatively urgent collection of ancillary data, or conduct of analyses to be reviewed at a further WG meeting in the near future).
- Report to the Director Research, DAFF BRANCH FISHERIES that exceptional circumstances exist and provide advice on the action to take.

The Chief Director Research, DAFF BRANCH FISHERIES will:

- Consider the advice from the WG.
- Decide on the action to take, or recommendations to make to his/her principals.

Examples of 'Principles for Action'

If the risk is to the resource, or to dependent or related components of the ecosystem, principles may be:

- The OMP-derived TAC should be an upper bound.
- Action should be at least an x% decrease in the TAC output by the OMP, depending on severity.

If the risk is to socio-economic opportunities within the fishery, principles may be:

- The OMP-derived TAC should be a minimum.
- Action should be at least a y% increase in the TAC output by the OMP, depending on severity.

For certain categories of exceptional circumstances, specific metarules may be developed and pre-agreed for implementation should the associated circumstances arise (for example, as has been the case for OMP's for the sardine-anchovy fishery where specific modified TAC algorithms come into play if abundance estimates from surveys fall below pre-specified thresholds). Where such development is possible, it is preferable that it be pursued.

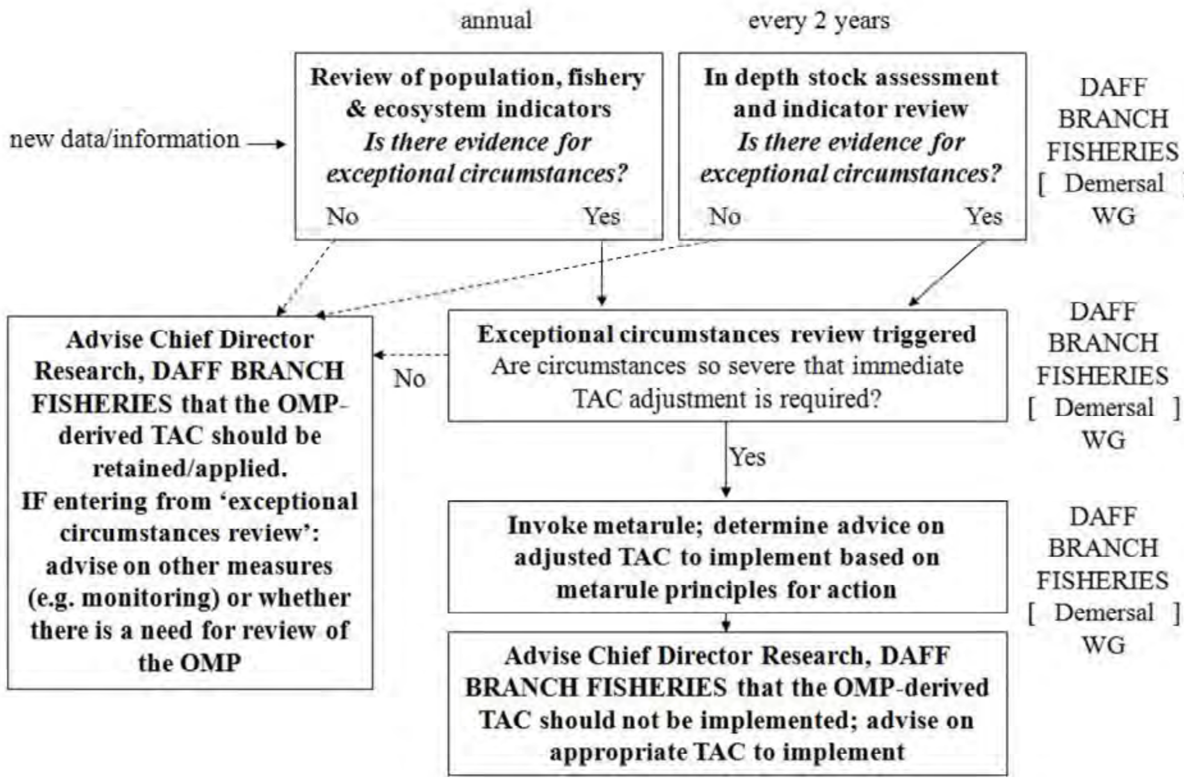


Figure D1: Flowchart for Metarules Process

D2. Regular OMP Review and Revision Process

The procedure for regular review and potential revision of the OMP is the process for updating and incorporating new data, new information and knowledge into the management procedure, including the operating models (OMs) used for testing the procedure. This process should happen on a relatively long time-scale to avoid jeopardising the performance of the OMP, but can be initiated at any time if the WG consider that there is sufficient reason for this, and that the effect of the revision would be substantial. During the revision process the OMP should still be used to generate TAC recommendations unless a metarule is invoked.

D2.1 Description of Process for Regular Review (see Fig.D2)

Every year the WG will:

- Consider whether the procedure for Metarule Process has triggered a review/revision of the OMP. Note that if proposals by a WG Member or Observer, or DAFF BRANCH FISHERIES Management, for an exceptional circumstances review include suggestions for an OMP review and possible revision, they must outline in writing the reasons why they consider this necessary, and must either indicate where the data or analyses are to be found supporting their proposed review, or must supply those data or analyses in advance of the WG meeting at which their proposal is to be considered. This includes the possibility of a suggested improvement in the manner in which the OMP calculates catch limitation recommendations; this would need to be motivated by reporting results for this amended OMP when subjected to the same set of trials as were used in the selection of the existing OMP, and arguing that improvements in anticipated performance were evident.

Every two years the WG will:

- Conduct an in depth stock assessment and review population, fishery and related ecosystem indicators, and any other relevant data or information on the population, fishery and ecosystem.
- On the basis of this, determine whether the assessment (or other) results are outside the ranges for which the OMP was tested (note that evaluation for exceptional circumstances would be carried out in parallel with this process; see procedures for the Metarule Process), and whether this is sufficient to trigger a review/revision of the OMP.
- Consider whether the procedure for the Metarule Process triggered a review / revision of the OMP.

Every four years since the last revision of the OMP the WG will:

- Review whether enough has been learnt to appreciably improve/change the operating models (OMs), or to improve the performance of the OMP, or to provide new advice on tuning level (chosen to aim to achieve management objectives).
- On the basis of this, determine whether the new information is sufficient to trigger a review/revision of the OMP.

In any year, IF the WG concludes that there is sufficient new information to trigger a review/revision of the OMP, the WG will:

- Outline the work plan and timeline (e.g. over a period of one year) envisaged for conducting a review.
- Report to the Chief Director Research, DAFF BRANCH FISHERIES that a review/revision of the OMP is required, giving details of the proposed work plan and timeline.
- Advise the Chief Director Research, DAFF BRANCH FISHERIES that the OMP can still be applied while the revision process is being completed (unless exceptional circumstances have been determined to apply and a metarule invoked).

In any year, IF the WG concludes that there is no need to commence a review/revision of the OMP, the WG will:

- Report to the Chief Director Research, DAFF BRANCH FISHERIES that a review/revision of the OMP is not yet required.

The Chief Director Research, DAFF BRANCH FISHERIES will:

- Review the report from the WG.
- Decide whether to initiate the review/revision process.

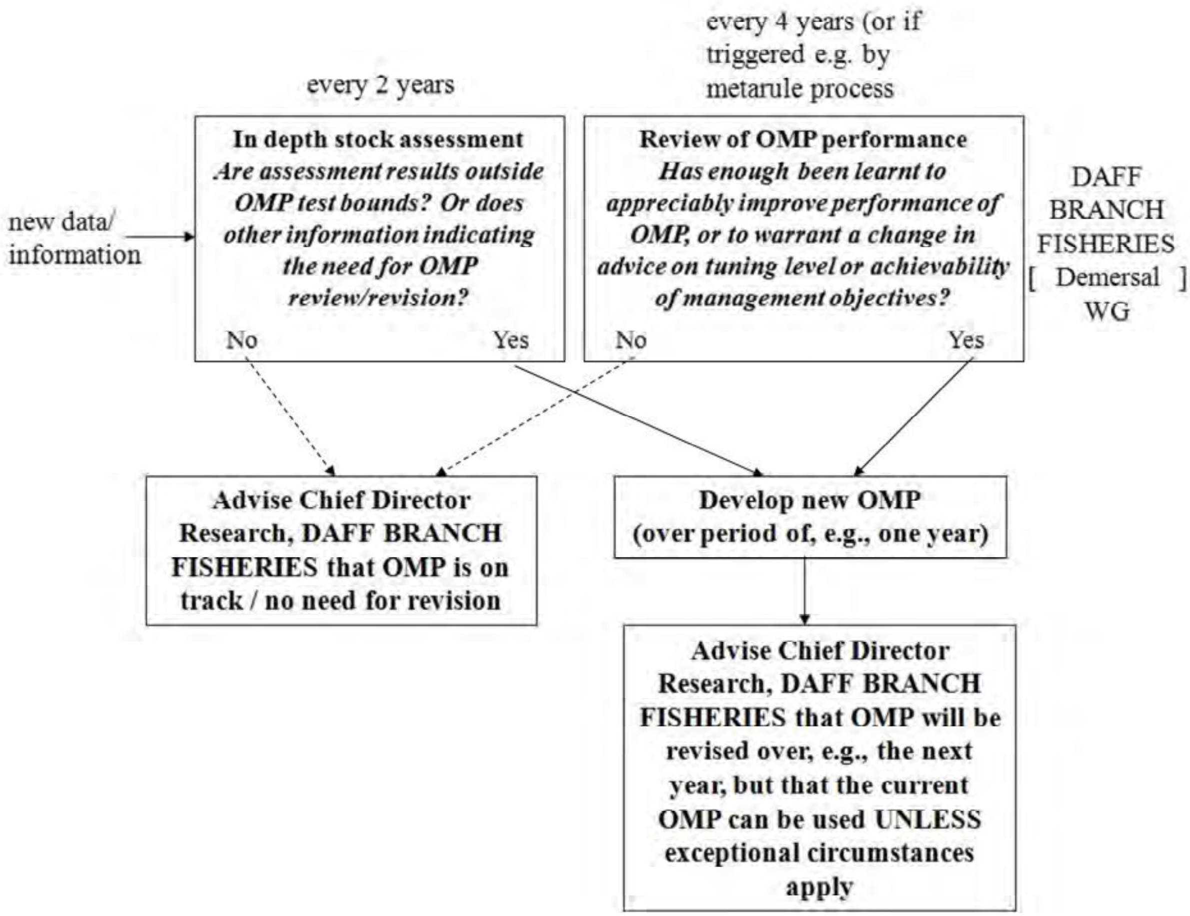


Figure D2: Flowchart for Regular Review and Revision Process

Appendix E

Projected future CPUE, survey abundance indices and recruitment

Figures E1-E2 plot the projected GLM-standardised CPUE and the survey abundance indices used in the OMP computations for each species for the RS under OMP-2018 respectively while Table E1 gives the 90% PI for each of these for the next four years. Note that the GLM-standardised CPUE series have been re-normalised by dividing by the 2016 value. This is done because the whole series changes when the GLM is rerun.

Figure E3 plots the projected proportion of *M. capensis* catch in the offshore trawl catch, with the 90% PIs for this proportion for the next four years are given in Table E3.

Table E1: 90% PI for the projected GLM-standardised CPUE and survey abundance indices (five-year running averages) for *M. paradoxus* and *M. capensis* for the RS under OMP-2018. Note: the new gear is assumed to be used on the *Africana* for all future surveys; if an industry vessel is used instead, the resultant estimates must be multiplied by 1.25 before comparison with the bounds in this table.

Year	West Coast CPUE ($CPUE_y/CPUE_{2016}$)	South Coast CPUE ($CPUE_y/CPUE_{2016}$)	West Coast summer survey	South Coast autumn survey
M. paradoxus				
2017	(0.76; 1.42)	(0.52; 1.43)		
2018	(0.71; 1.60)	(0.50; 1.63)	(123.1; 707.4)	(21.8; 157.0)
2019	(0.67; 1.81)	(0.49; 1.87)	(122.9; 772.0)	(21.5; 175.0)
2020	(0.65; 1.96)	(0.48; 2.04)	(124.7; 808.0)	(21.6; 191.9)
2021			(127.7; 831.6)	(22.1; 203.7)
M. capensis				
2017	(0.82; 1.58)	(1.03; 2.52)		
2018	(0.84; 1.79)	(1.07; 2.82)	(49.0; 236.0)	(60.4; 252.7)
2019	(0.86; 1.97)	(1.09; 3.08)	(48.1; 244.5)	(58.9; 263.3)
2020	(0.86; 2.15)	(1.10; 3.30)	(47.5; 250.8)	(56.8; 272.4)
2021			(47.5; 259.3)	(55.3; 282.2)

Table E2: 90% PI for the projected proportion of *M. capensis* in the offshore trawl catch.

Year	West Coast	South Coast
2018	(0.07; 0.27)	(0.04; 0.29)
2019	(0.07; 0.27)	(0.04; 0.30)
2020	(0.07; 0.27)	(0.04; 0.30)
2021	(0.07; 0.28)	(0.04; 0.30)
2022	(0.07; 0.28)	(0.04; 0.31)

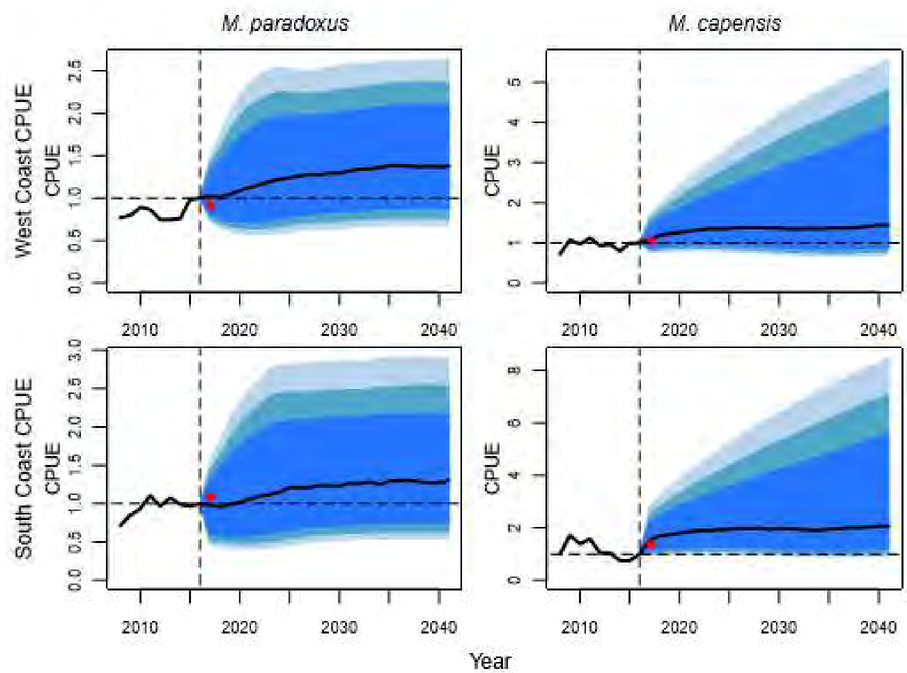


Fig. E1: 95, 90, 80% PE and median for the projected GLM-standardised CPUE for *M. paradoxus* and *M. capensis* for the RS under OMP-2018. The red dots show the 2017 CPUE indices, standardised relative to the 2016 value in the updated GLM series.

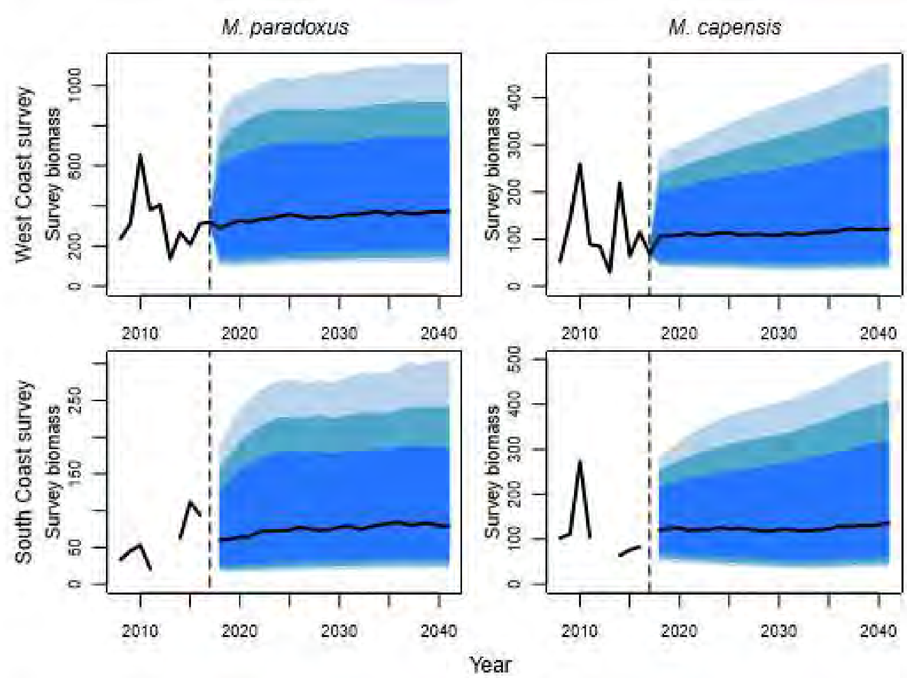


Fig. E2: 95, 90, 80% PE and median for the survey abundance indices for *M. paradoxus* and *M. capensis* for the RS under OMP-2018. Gaps in the median trajectory for the South Coast survey indicate surveys that did not take place. Since no surveys took place in 2018, no further data have been added to the projection PEs. Note: future surveys are assumed to be carried out using the new gear on the *Africana*; if an industry vessel is used instead, the resultant estimates must be multiplied by 1.25 before comparison with the bounds in these plots.

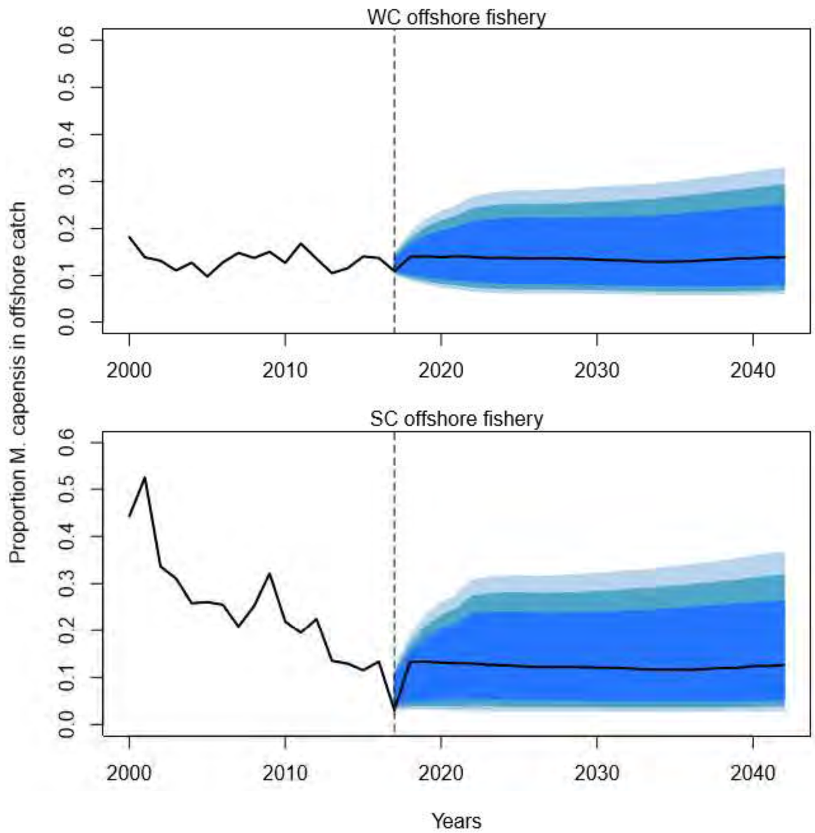


Fig. E3: 95, 90, 80% PE and median for the proportion *M. capensis* in the offshore trawl catch.

Gulf States Marine Fisheries Commission
2404 Government Street
Ocean Springs, Mississippi, 39564

www.gsmfc.org