

Two classes of multi-model candidate management procedures for Atlantic bluefin tuna.

Sean P. Cox, Samuel D. N. Johnson, Stephen P. Rossi,

Abstract

We developed two classes of management procedures for Atlantic bluefin tuna based on multi-model inference. The basis of the procedures were estimation methods tuned to five operating models from the reference OM grid. For the empirical class, we used OM catchability and a stationary stock mixing distribution to estimate area biomass from the larval indices. For model-based class of MPs, we tuned 5 delay difference assessment models to each of the 5 operating models, matching stock recruit steepness and biomass for the recent historical period from 1965 - 2016. At each time step, estimates of current (empirical) or projected (model-based) biomass were generated from approved management indices. These estimates were used in harvest control rules to generate area-specific TACs, and the five TACs were averaged to produce harvest advice for the East and West area. Multiple MPs were then defined based on varying precautionary TAC caps, maximum target harvest rates, and HCR control points. We found that MPs with lower caps, lower maximum harvest rates and control points avoided overfishing on the reference grid more often. We also found that the subset of OMs that our AMs were tuned to managed to capture the uncertainty well.

September 19, 2019

Contents

| | | |
|----------|---|-----------|
| 1 | Introduction | 2 |
| 2 | Methods | 3 |
| 2.1 | Tuning subset of operating models | 3 |
| 2.2 | Empirical estimation procedure | 3 |
| 2.3 | Delay difference model | 4 |
| 2.4 | Harvest control rules | 6 |
| 2.5 | Providing catch advice by area | 7 |
| 2.6 | Simulation experiments and performance metrics | 7 |
| 3 | Results | 8 |
| 3.1 | Fits of Delay Difference AMs to historical data | 8 |
| 3.2 | Performance in projections | 9 |
| 4 | Discussion | 10 |
| 5 | Tables | 12 |
| 6 | Figures | 16 |

1 Introduction

The International Commission for the Conservation of Atlantic Tunas (ICCAT) is currently conducting a management strategy evaluation (MSE) for Atlantic bluefin tuna (ref ICCAT 2019). The technical core of an MSE is closed loop simulation testing of candidate management procedures (CMPs) that provide harvest advice in the presence of uncertainty. Uncertainty is captured by a set of operating models, which represent a plausible range of states for the management system.

The operating models for Atlantic bluefin tuna are based on different hypotheses (fits) of a multi-stock, multi-area, age- and length-structured model known as the Modifiable Multi-stock Model (M3), with 4 fishing seasons per year, that allow for mixing of fish from both stocks in each area in a given season (**REF TSD**). Given the inherent complexity of such a model, it is impossible to use the same structure within a management procedure, as management procedures must be able to be tested within simulations in a timely manner. Therefore, management procedures must be designed that are of lower levels of complexity than the operating model.

In this paper, we develop and test two classes of CMPs that use multi-model inference to provide harvest advice in the MSE framework. The first class is based on an empirical index-based method of estimating spawning stock biomass, and the second replaces the empirical method with a multi-stock state-space delay difference stock assessment model. Both classes are made up of five biomass estimation and harvest control rules that are tuned to one of five reference grid operating models. Catch advice from each sub-procedure is then combined to produce area-specific TACs.

2 Methods

Our candidate management procedures were based on multi-model inference [rossi2019inferring], combining estimates of biomass and associated catch advice from sub-procedures tuned to a subset of the reference grid of multi-stock operating models. This multi-model formulation allowed us to capture the biological uncertainty represented by the reference grid of operating models.

We defined two classes of MP: empirical and model based. The empirical MPs scale larval indices by OM estimates of catchability, while the model-based MPs fit state space delay difference assessment model to management indices.

2.1 Tuning subset of operating models

We chose a tuning subset of five reference grid operating models, namely OMs 1, 2, 4, 7, and 11, to tune our estimation procedures. These five OMs were chosen to span the three main axes of operating model uncertainties (stock-recruit steepness, natural mortality/maturity, and stock mixing), as well as the range of spawning stock biomass scales and dynamics (Figure 1).

2.2 Empirical estimation procedure

The empirical estimation procedure used a simple moving average of the two spawning stock larval survey indices. For each time step t , and stock $s \in \{GOM, MED\}$ the current average larval index is calculated as

$$\bar{I}_{s,t} = \frac{\sum_{t'=t-3}^t I_{s,t'}}{4}, \quad (1)$$

where $I_{s,t}$ is the larval index for stock s at time t . We then scaled the smoothed survey indices to multiple biomass estimates, one for every operating model in the tuning subset, by using the appropriate larval survey catchability parameter, e.g

$$\hat{B}_{s,t,k} = \frac{\bar{I}_{s,t}}{q_{s,k}}, \quad (2)$$

where $q_{s,k}$ is the stock larval index catchability parameter from the tuning operating model $k \in \{1, 2, 4, 7, 11\}$.

Spawning stock biomass was then translated to area biomass using an assumed constant distribution of biomass for stock mixing

$$P = (p_{s,a}) = \begin{pmatrix} .898 & .102 \\ .1 & .9 \end{pmatrix}, \quad (3)$$

found by averaging the proportion of stock biomass in each area over the historical period in the tuning set of OMs. The rows of matrix P are indexed by spawning stocks $\{MED, GOM\}$, and the columns are indexed by management area $\{E, W\}$. This generates an area specific biomass

$$B_a = \sum_s p_{s,a} B_s, \quad (4)$$

where $s \in \{MED, GOM\}$, $a \in \{E, W\}$.

2.3 Delay difference model

For the model-based class of procedures, the simple index scaling method of the empirical procedure is replaced by a state-space multi-stock Schnute-Deriso delay difference assessment model [@schnute1985general; @deriso1980harvesting]. The two spawning stocks are assumed to be distinct spawning stocks for the larval surveys, but are mixed using the distribution P for the other management indices and the observed catches.

2.3.1 Equilibria

Because there is no mixing of spawners, each stock has the simple Delay Difference equilibrium states, which can be expressed as a function of long-term fishing mortality $f = F_s$. The equations for equilibria in Table 1 were used to initialise the model at a fished equilibrium in 1965, as well as estimate biological reference points for use in the harvest control rules.

2.3.2 Biomass time series reconstruction

Each assessment model was initialised in a fished state in 1965. To do so, we estimated an initial fishing mortality rate $F(s, 0)$, and assumed that the stock was at the fished equilibrium defined in Table 1. Other free parameters were unfished biomass $B_{s,0}$, catchability q_g and observation error uncertainty τ_g for biomass and abundance indices, and recruitment process error deviations $\omega_{s,t}$ for $t \in 1966, \dots, 2014$ (Table 2). Stock recruit steepness h_s and natural mortality M_s were fixed to the estimated values from the associated operating model.

We assumed the that proportion $p_{s,a}$ of stock $s \in \{MED, GOM\}$ in area $a \in \{E, W\}$ was constant, allowing the fluctuations in stock biomass to account for the mixing dynamics. We also assumed that the stocks were homogeneously mixed in each area, so that total stock specific catch was the sum of the area specific catches, split according to the proportion to the stock biomass in each area (eqn ref).

We used a numerical Newton-Rhapon method to solve for instantaneous fishing mortality at all time steps $t \geq 1$, which were used in the survival calculation to progress numbers and biomass to the next time step.

We modeled recruitment process errors as a simple random walk, rather than indepent deviations from the stock recruit curve. We chose a random walk as it was more able to capture dynamic recruitment regimes. Furthermore, random walks would also make projected recruitments, which are an important component of projected biomass in a delay difference formulation, less likely to be average, and more like the estimated recruitment in the last year of the assessment period.

Because this is a model-based MP and the time-series of approved management indices are rather short, we included two additional indices. These indices were the M3 model yearly spawning stock biomass estimates for the East and West stocks. These biomass time series were assumed to be observed with catchability $q = 1$ and an observation error CV of 5%. This allowed our AMs to estimate catchability for the approved indices and scale them appropriately to the spawning stock biomass from the associated operating model, which was required for extending the fit to the approved indices in the projections, as well as making sensible estimates of biological parameters despite the complexity mismatch between the operating model and the assesment models.

Finally, we applied a log-normal prior distribution to unfished biomass $B_{0,s}$. This was applied to prevent the unfished biomass from being estimated too close to the initial biomass for the assessment period, and producing optimistic estimates of initial and current biomass depletion. For each AM, the prior mean was defined as the estimate from the associated operating model and we assumed a log standard deviation of 0.05.

2.3.3 Reference Points

Reference points for each stock were estimated from the equilibrium values in (Table 1). The optimal fishing mortality rate F_{MSY} was found by numerically solving for the stationary point of the yield curve, and was in turn used to estimate the optimal equilibrium biomass B_{MSY} associated with that mortality rate.

We estimated area based reference points for use in harvest control rules, which were the stock-area

biomass weighted averages of the stock-specific reference points, i.e.

$$B_{MSY,a} = \sum_s \frac{B_{s,a}}{B_a} B_{MSY,s}, \quad (5)$$

$$F_{MSY,a} = \sum_s \frac{B_{s,a}}{B_a} F_{MSY,s}, \quad (6)$$

where $B_{s,a} = p_{s,a} B_s$.

2.4 Harvest control rules

For each of the five sub-procedures we defined a ramped harvest control rule with upper and lower control points and a maximum harvest rate. The control points and maximum harvest rates for these rules were based either based on the associated operating model's biological reference points, or the , and a precautionary TAC cap was used to limit removals in cases of bias in projected biomass or reference point estimates.

Our ramped harvest control rule was based off the rules used for albacore tuna [ref]. These rules requires an upper control point (UCP), a lower control point (LCP) and a maximum target harvest rate U_{max} . The general form of the rules (ignoring AM, area, and stock indices):

$$U_{targ} = \begin{cases} 0.1 \cdot U_{max} & B \leq LCP \\ U_{max} * (0.1 + 0.9 \frac{B}{UCP-LCP}) & LCP \leq B \leq UCP \\ U_{max} & UCP \leq B \end{cases} \quad (7)$$

For all HCRs, we set $LCP = .4UCP$, and tested two options for the upper control point: $UCP = B_{MSY}$ or $UCP = .4B_0$. We also tested two options for the maximum target harvest rate at $U_{max} = F_{MSY}$ and $U_{max} = \frac{2 \cdot M}{3}$. Under the empirical management procedures, control points were taken from the tuning grid of operating models, while under the model-based MPs, control points were taken as the delay difference model equilibria.

We applied a TAC cap in each area to make the HCRs more precautionary. This extra precaution is required to guard against optimistic assessment errors and biases in reference point calculations, both of which are

caused by the complexity mismatch between the AMs and the OMs. We tested three kinds of caps: a high cap of 25 kt in the East, and 4 kt in the West; a low cap of 20 kt in the East, and 2.5 kt in West; and a cap of the stock and area based MSY, which was derived from assessment model equilibria (Tables 1 and 3, Figure 2).

Both home-stock- and area-based harvest control rules were applied for East and West area TACs. For example, in the East area, HCRs were applied based on East (MED) spawning stock biomass compared to the East stock control points and harvest rates, and East area biomass compared to mixed stock control points and harvest rates. Mixed area based control points and harvest rates were averaged over the two stocks present in the area, weighted by the proportion of stock specific biomass in that area (e.g. see area-based reference points calcs in previous section), and the lower of the two TACs was chosen. From this, the West TAC would almost always be managed according to the Gulf of Mexico spawning stock harvest control rules, and the East TAC would almost always be managed according to the mixed East Area harvest control rule, as this includes the weaker stock.

2.5 Providing catch advice by area

To provide a single TAC at each time step we averaged over the TAC advice from the 5 sub-procedures for both the empirical and model-based MPs. For the model-based MP, we tested two weightings for the averaging: equal weighting (standard mean), and weighting by Akaike’s Information Criterion (AIC). For the empirical MPs, we fully tested only a simple average weighting.

To calculate the AIC weighting across the tuning grid, we used only the observation model likelihood components for the approved management indices. This is because the approved management indices were data that were common among the tuning grid AMs, while the included OM spawning stock biomass indices differed between AMs, invalidating the assumptions of AIC.

Final TACs were then smoothed with respect to the previous management interval’s TAC. This smoothing allowed a maximum increase of 20%, and a maximum decrease of 50%, compared to the previous management interval.

2.6 Simulation experiments and performance metrics

Rather than run the entire grid of management procedures defined by Table 3, we ran three batches of CMPs:

1. Empirical MPs with MSY caps:

- **empMP_msyCap, empMP_msyCap.4B0**

2. Equal weighted delay-difference MPs with MSY caps:

- **MP_msyCap, MP_msyCapF23M, MP_msyCapF23M.4B0**

3. AIC weighted delay difference MPs with MSY caps:

- **MP_aic_msyCap, MP_aic_msyCapF23M, MP_aic_msyCapF23M.4B0**

The suffix *.4B0* indicates that the proxy UCP of *.4B₀* is applied, and F23M indicates that the maximum harvest rate is set at $\frac{2}{3}M$, where M is fixed to the associated operating model in the tuning subset.

All MPs were evaluated over the 12 reference grid operating models, labelled OM_1 through OM_12, as well as three additional robustness operating models, OM_13 to OM_15. To compare performance of each MP, we calculated distributions of the performance metrics

$$B_{30} = \frac{B_{2046}}{B_{MSY}}, \quad (8)$$

$$AvC_{30} = \frac{\sum_{t=2016}^{2046} C_t}{30} \quad (9)$$

over the set of simulation replicates for each stock, MP and operating model combination.

3 Results

3.1 Fits of Delay Difference AMs to historical data

We found that by fitting the delay difference assessment model to operating model biomass, we could recapture the operating model biomass series faithfully (Figures 2 and 3). Management indices were then able to scale to the stock biomass and estimates of catchability parameters were suitable for the delay difference assessment (Figure 3).

We found that assessment model estimates of biological parameters and equilibria, and operating model estimates of the same quantities, were less consistently similar. While the log-normal prior on unfished biomass worked well to keep $B_{0,s}$ estimates close to the operating model values, the model equilibria were often biased. For example, in AMs 1, 4, and 7, the DD model F_{MSY} values for the Mediterranean spawning stock are at least 0.17, in contrast with the operating model values, which were at most 0.13. Similarly, in the same assessment models, B_{MSY} values were around 50% of the corresponding operating model values. This led to similar values for maximum sustainable yield MSY , indicating that the biases in F_{MSY} and B_{MSY} were negatively correlated. These biases have implications for the performance of MPs that rely on the AM reference point estimates to set TACs, which we describe below.

3.2 Performance in projections

3.2.1 Empirical MPs

As expected, empirical CMPs did a decent job of predicting future spawning stock biomass (Figure 6). This is unsurprising, as empirical CMPs were tightly tuned to the tuning subset of OMs, using catchability from the two larval surveys, and OM reference points for the harvest control rules in each sub-procedure.

We found that the proxy upper control point was not necessary for the empirical MPs. This was indicated by the almost identical performance metrics between those MPs that used OM B_{MSY} , and those that used 40% of OM B_0 , as the harvest control rule's upper control point (Figure 7).

Empirical MP performance was more variable across reference grid operating models for the East stock than in the West. For example, under MPs that used the OM MSY as the TAC cap, B_{30} ranged between 0.5 and 1.5 in the East for all reference and robustness OMs, while in the west the same metric was often between 1.5 and 2 on the reference grid, with the exception of OMs 4 and 5.

The high values of B_{30} in the west probably indicate that the empirical MPs are too conservative for that stock, and TACs could be set higher for fishing that catches the Gulf of Mexico stock.

There was very little difference in the AvC_{30} metric between operating models in the East, indicating that the MPs tended to set TACs around the caps. In the west, there was more variability in the distributions of AvC_{30} , but the medians were similar across OMs.

3.2.2 Delay Difference Model-based MPs

Similar to the empirical MPs, we found the model based MPs to fit the spawning stock biomass in a tuning subset operating model reasonably well (Figure 8). Again, this is unsurprising given the close tuning each AM has to its associated tuning subset operating mode.

We found MPs that used proxy values for the upper control point and maximum target harvest rates produced higher ratios of biomass to B_{MSY} in 2046 than those that used the delay difference model reference points (Figures 9 and 10, upper panels). This was due to the bias in the delay difference reference points discussed above, where F_{MSY} was positively biased and B_{MSY} was negatively biased. Because of this, although some AM estimates of current and projected biomasses were often very close to the OM biomass, the biomass relative to B_{MSY} was above the upper control point, so the positively biased maximum target harvest rate was applied and overfishing occurred.

The effects of the proxy harvest rate and upper control point were different in the east and the west. In the east, the two proxy MPs we tested for both mean and AIC weighted TACs produced very similar B_{30} and AvC_{30} distributions across all OMs (Figures 9 and 10, left hand panels). In contrast, in the West

the B_{30} metric was most affected by the **msyCapF23M.4B0** MP with both proxies applied, while the **msyCapF23M** MP had less of an effect. The difference was more evident in the average catch metric, with the **msyCapF23M.4B0** MP reducing average catch by about 500t, compared to an often negligible reduction for the **msyCapF23M** (Figures 9 and 10, right hand panels).

The AIC weighted MPs (Figure 9) were a little less precautionary than the equal weighted MPs (Figure 10). Average catches were about 10% higher, and relative biomass a little lower. In general, the relative ranking of MPs with the AIC and equal weighting classes was the same, as discussed above; however, the distance between the medians values of each metric within an OM was a little greater for the AIC weighted class. This is likely an indication of the shifting TAC weighting at each time step, which allows the TAC to change more rapidly as new data becomes available.

3.2.3 Performance of CMPs on OMs outside of the tuning subset

Our CMPs had mixed performance on the reference grid OMs outside of the tuning subset. These were OMs 3, 5, 6, 8, 9, 10 and 12.

Overall, our MPs had low B_{30} values on OMs 3, 6, 9, and 12 in the East; however, these OMs have a recruitment regime shift that was erroneously implemented (Carruthers, pers. comm., 2019), switching the recruitment regimes between spawning stocks and time periods, creating a larger reduction in steepness values between regimes than was intended in the East, and a smaller reduction in the West.

On the remaining OMs, that differ from the tuning subset by mortality/maturity and stock mixing, the AIC weighted model-based MPs performed the best on the B_{30} metric. Out of those, the AIC weighted MP **MP_aic_msyCapF23M** performed the best in the East, coming the closest to $B_{30} = 1$ on OMs 5, 8, 9, and 10 and 12. The same MP was more precautionary in the West, often selecting lower TACs than necessary under the AIC weighting (Figure 11).

4 Discussion

In this paper, we tested multiple empirical and model-based candidate management procedures (CMPs) for Atlantic Bluefin Tuna. These CMPs were based on multi-model inference, where TACs from 5 procedures, each tuned to a specific reference grid OM, were combined in some way to produce TAC advice. For empirical CMPs, we used an equal weighting, while for model based delay-difference CMPs, we tested an equal weighting as well as an AIC-like weighting based on data shared between the tuned AMs.

We found tuning the CMPs to perform similarly across the OMs and stocks was a challenge. Our MPs were on average too precautionary in the West area/stock, and not precautionary enough on the East

area/stock. This could be avoided by using combinations of MPs with different caps or harvest rates in the East and West, allowing the West catch to increase while keeping the East stable.

Similarly, choosing the correct weighting to combine the TAC advice was challenging. The equal weighting scheme is more robust to the issues of the AIC weighting (small differences in AIC values tend to downweight models disproportionately), but led to MPs that were often too precautionary. On the other hand, while the AIC weighting was often biased, it did allow the weights to shift between assessment periods, which created both over and under harvesting on occasion.

Extending the empirical MP to use an unequal weighting is our next step. We conceived of a stochastic approximation algorithm to optimise TAC weightings [spall2005introduction]. This algorithm would apply gradient descent to an objective function based on median B_{30} and ∇B_{30} values for a set of operating models to TAC weightings. Preliminary tests show promise, but this is still very much under development.

5 Tables

Table 1: Unfished and fished equilibrium quantities for the delay-difference population dynamics.

| Description | Equation |
|----------------------|--|
| Survivorship | $S^{(f)} = e^{-M-f}$ |
| Average Weight | $\bar{w}^{(f)} = \frac{S^{(f)} \alpha + w^{(k)} (1 - S^{(f)})}{1 - \rho S^{(f)}}$ |
| Unfished Numbers | $N = B_0 / \bar{w}^{(f=0)}$ |
| Unfished Recruitment | $R_0 = (1 - S^{(f=0)}) N_0$ |
| Stock-Recruit | $a = \frac{4hR_0}{B_0(1-h)}, b = \frac{5h-1}{B_0(1-h)}$ |
| Biomass | $B^{(f)} = \frac{S^{(f)} (\alpha + \rho \bar{w}^{(f)}) + \bar{w}^{(f)} (a w^{(k)} - 1)}{b(\bar{w}^{(f)} - \rho S^{(f)} \bar{w}^{(f)}) - \alpha S^{(f)}}$ |
| Recruitment | $R^{(f)} = \frac{a B^{(f)}}{1 + b B^{(f)}}$ |
| Yield | $Y^{(f)} = \frac{f}{M+f} (1 - e^{-M-f}) B^{(f)}$ |

Table 2: Process and observation model components of the delay difference stock assessment model used in BC Sablefish management procedures. Initialisation values for biomass, numbers, and recruitment are equilibrium unfished values from Table 2.

| No. | Equation |
|--|---|
| Model Free Parameters | |
| A2.1 | $\theta_s = \{\log F_{s,0}, \log B_{s,0}, q_g, \tau_g, \vec{\omega}_{s,t}\}.$ |
| Initialisation ($t = 1$) | |
| A2.2 | $B_{s,1} = B_s^{(f=F_{s,0})}$ |
| A2.3 | $N_{s,1} = N_s^{(f=F_{s,0})}$ |
| A2.4 | $R_{s,1} = R_s^{(f=F_{s,0})}$ |
| Mixed Area Biomass | |
| A2.5 | $B_{s,a,t} = p_{s,a} \cdot B_{s,t}$ |
| A2.6 | $B_{a,t} = \sum_s B_{s,a,t}$ |
| Stock Specific Catch | |
| A2.7 | $C_{s,t} = \sum_a \frac{B_{s,a,t}}{B_{a,t}} C_{a,t}$ |
| Time series Reconstruction ($2 \leq t \leq T$) | |
| A2.8 | $R_{s,t} = \begin{cases} R_{s,1} e^{\omega_{s,t-k} - \sigma^2/2} & t \leq k \\ \frac{aB_{s,t-k}}{1+bB_{s,t-k}} e^{\omega_{s,t-k} - \sigma^2/2} & t > k \end{cases}$ |
| A2.9 | $Z_{s,t-1} = M_s + F_{s,t-1}$ |
| A2.10 | $N_{s,t} = e^{-Z_{s,t-1}} N_{s,t-1} + R_{s,t}$ |
| A2.11 | $B_{s,t} = e^{-Z_{s,t-1}} (\alpha N_{s,t-1} + \rho B_{s,t-1}) + w_s^{(k)} R_{s,t}$ |
| Statistical Model | |
| A2.12 | $\hat{I}_{g,t} = q_g \cdot B_{s,t}$ |
| A2.13 | $\hat{I}_{g,t} = q_g \cdot B_{a,t}$ |
| A2.14 | $\log I_{g,t} \sim N(\log \hat{I}_{g,t}, \tau_g)$ |
| A2.15 | $\omega_{s,t} \sim N(0, \sigma)$ |

Table 3: Harvest control rule settings that we tested. A harvest control rule is defined by taking one element from each column.

| Upper Control Point | Maximum harvest rate | TAC Caps |
|---------------------|----------------------|------------------|
| B_{MSY} | F_{MSY} | (MSY, MSY) |
| $.4B_0$ | $\frac{2}{3}M$ | $hi = (4, 25)$ |
| | | $lo = (2.5, 20)$ |

Table 4: Delay difference assessment model estimates of unfished biomass, natural mortality, and biological reference points, when fit to the tuning grid of operating models 1, 2, 4, 7, and 11.

| AM | Mediterranean Spawners | | | | | Gulf of Mexico Spawners | | | | |
|----|------------------------|---------------|---------------|-------------|-----------|-------------------------|---------------|---------------|-------------|-----------|
| | $B_{0,MED}$ | $B_{MSY,MED}$ | $F_{MSY,MED}$ | MSY_{MED} | M_{MED} | $B_{0,GOM}$ | $B_{MSY,GOM}$ | $F_{MSY,GOM}$ | MSY_{GOM} | M_{GOM} |
| 1 | 1503.86 | 243.93 | 0.24 | 58.77 | 0.10 | 123.08 | 26.28 | 0.17 | 4.60 | 0.10 |
| 2 | 1578.01 | 432.77 | 0.09 | 39.93 | 0.10 | 186.78 | 57.32 | 0.07 | 4.28 | 0.10 |
| 4 | 1728.01 | 283.50 | 0.17 | 47.10 | 0.07 | 108.41 | 32.49 | 0.05 | 1.76 | 0.07 |
| 7 | 1532.64 | 248.60 | 0.24 | 59.89 | 0.10 | 182.32 | 38.93 | 0.17 | 6.81 | 0.10 |
| 11 | 2208.68 | 590.72 | 0.07 | 39.53 | 0.07 | 384.48 | 115.22 | 0.05 | 6.23 | 0.07 |

6 Figures

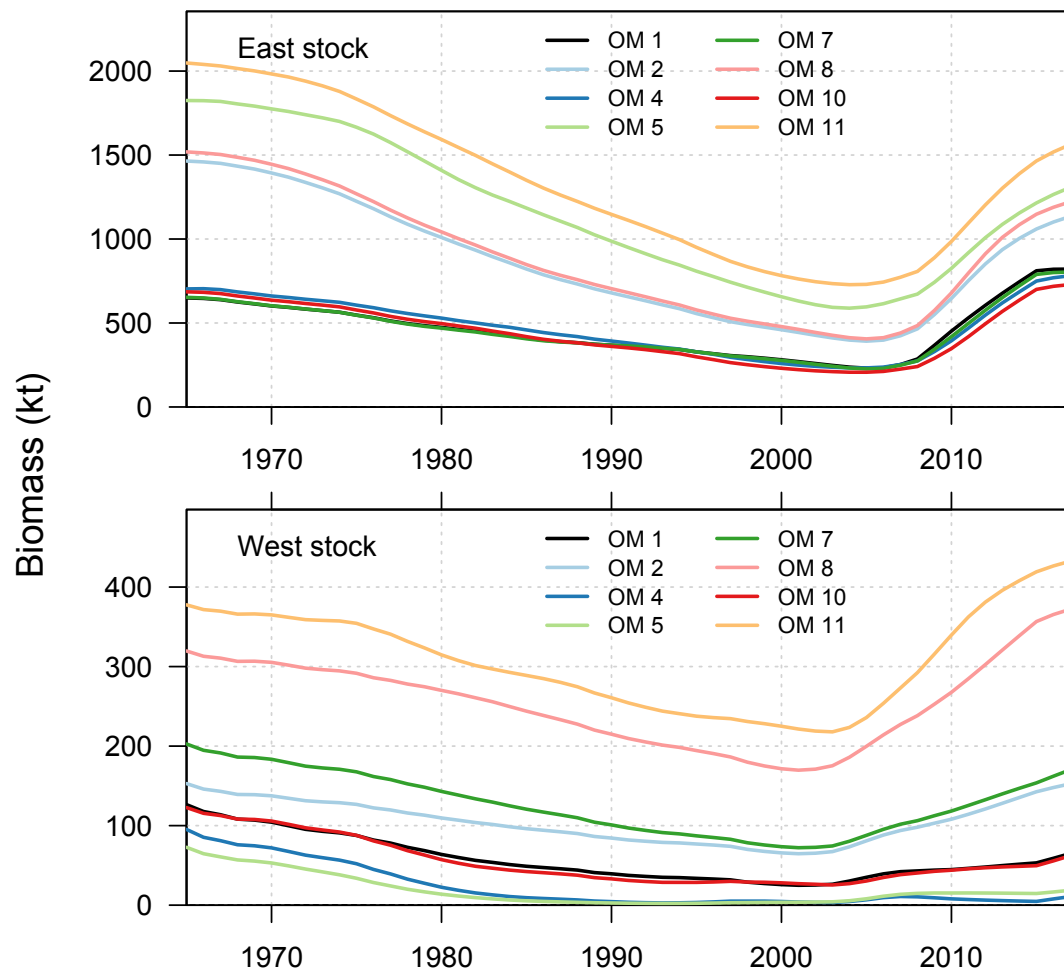


Figure 1: Operating model spawning stock biomasses for the full reference grid for both the Mediterranean stock (East) and Gulf of Mexico stock (West). Each colour corresponds to a different OM from the reference grid.

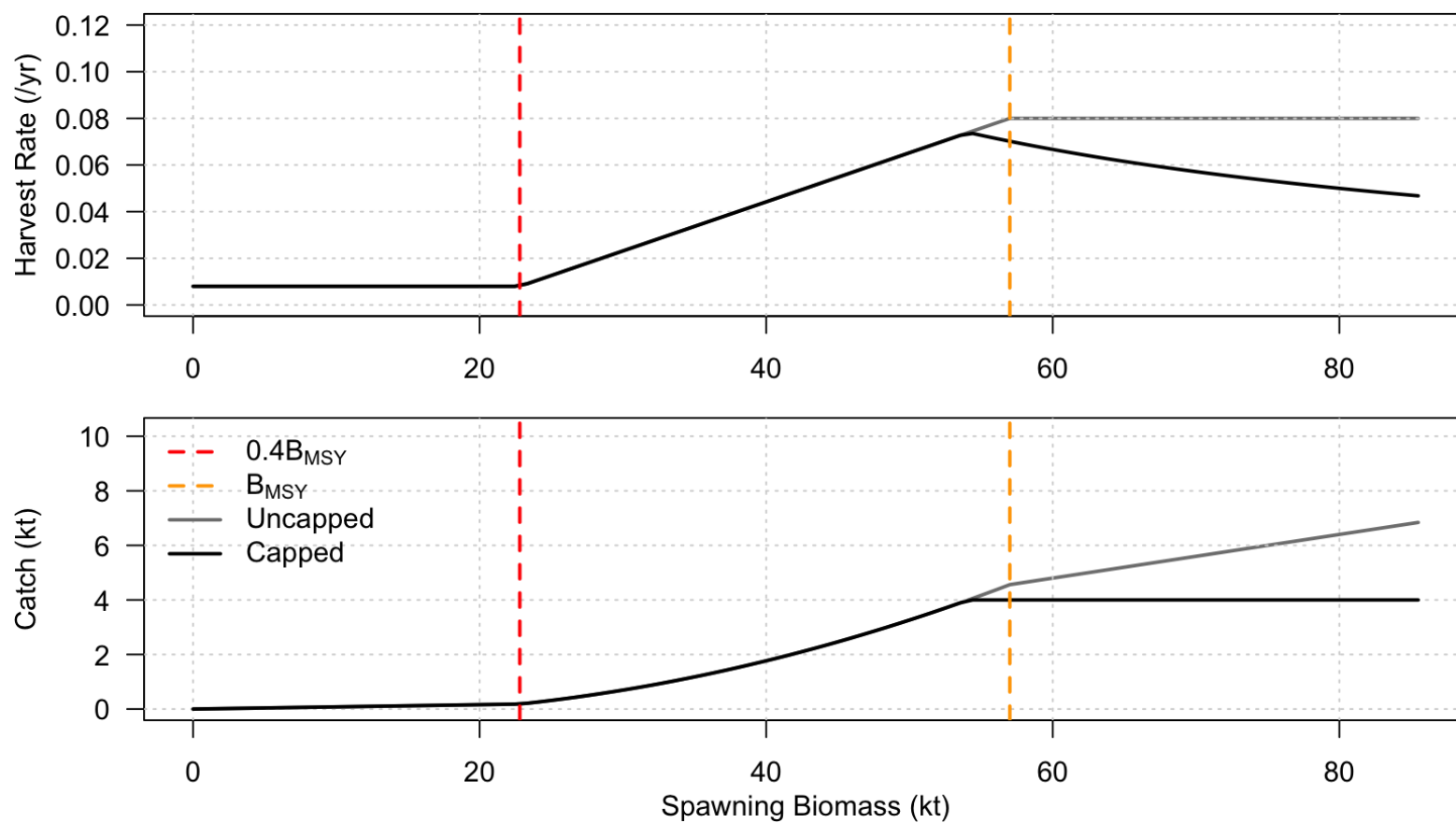


Figure 2: An example harvest control rule with maximum harvest rate $U_{max} = 0.08$, an upper control point of $B_{MSY} = 57$, and a TAC cap at 4 kt.

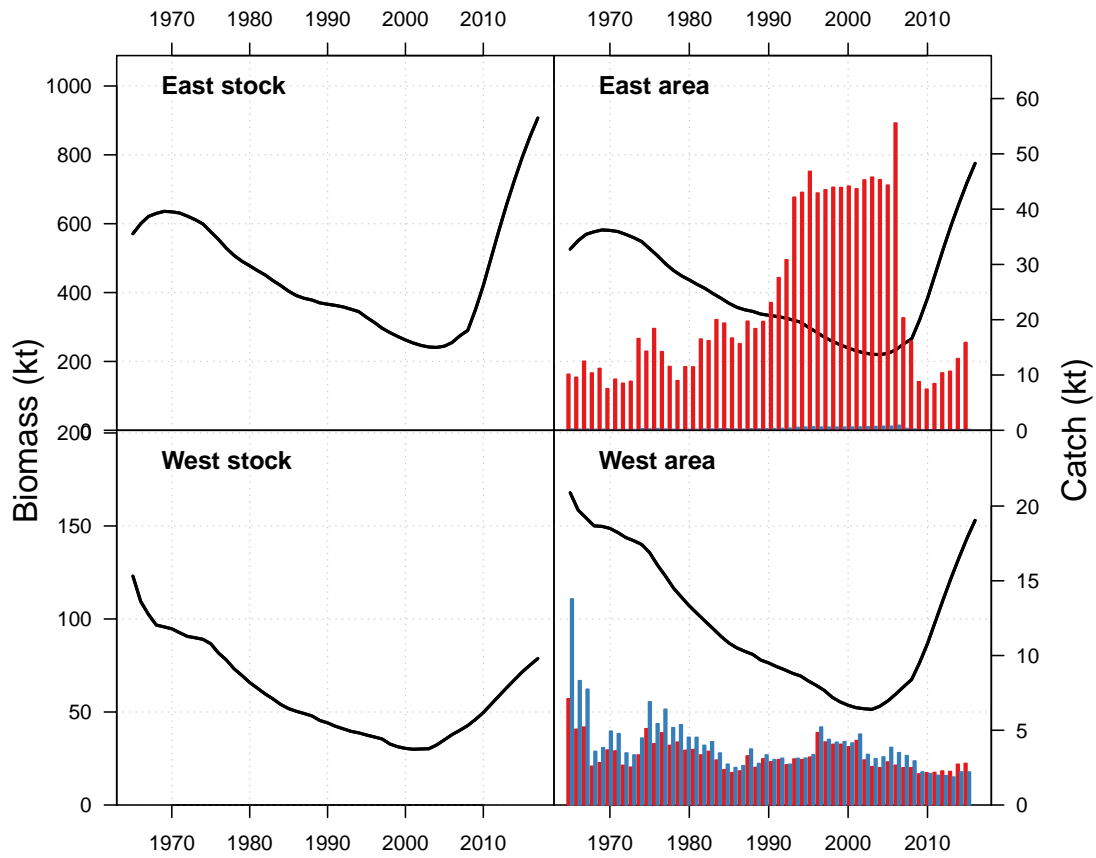


Figure 3: Spawning stock biomass estimates by stock (left hand column) and area (right hand column) from the delay difference stock assessment model fit to OM 1. Total area catch is shown split into East stock (red bars) and West stock (blue bars). Note that the catch scale is exaggerated with respect to the biomass, so that West stock catch is visible in the East area.

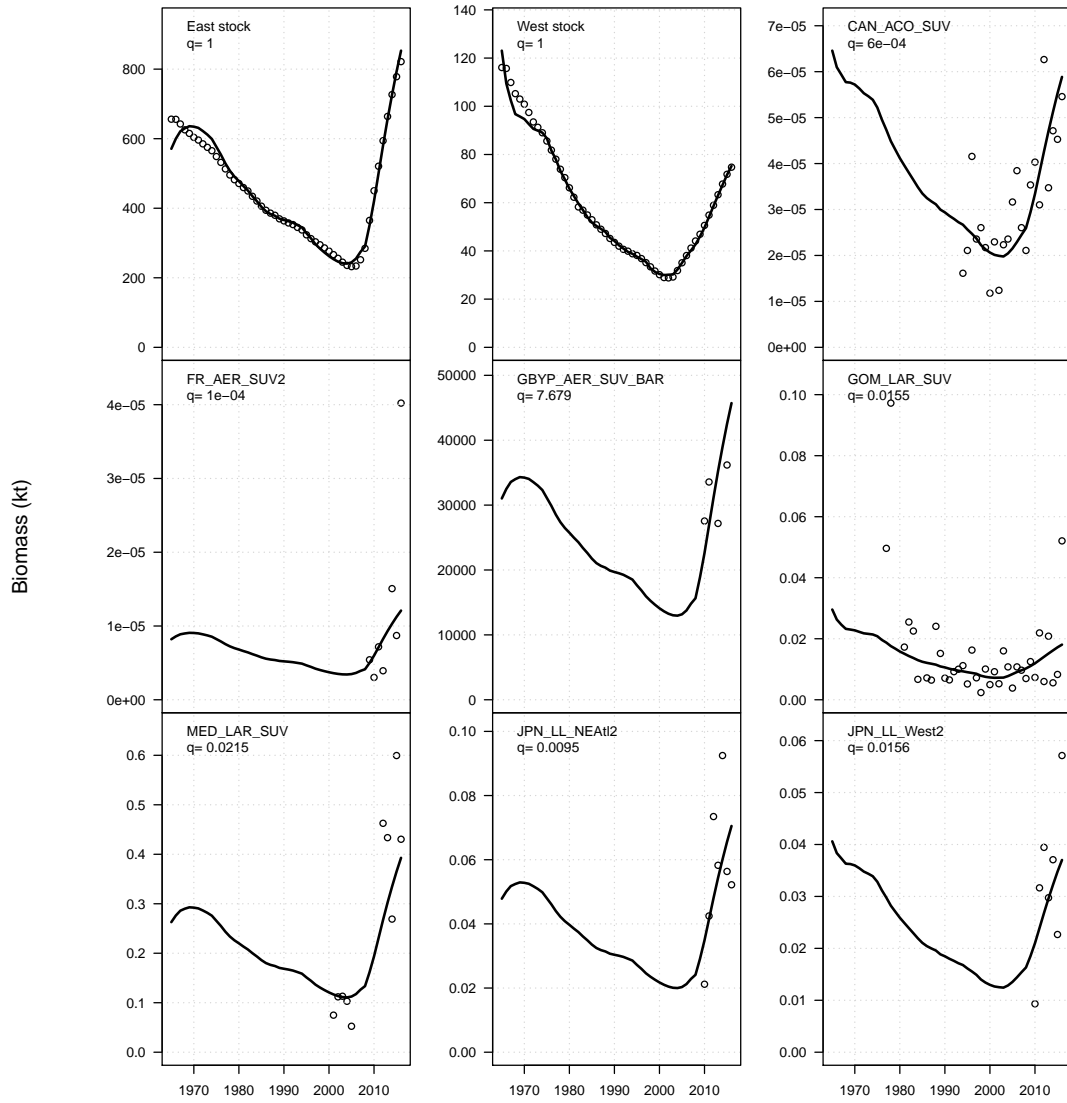


Figure 4: Fits of the delay difference stock assessment model to stock and area management indices, with the associated catchability estimates. Data are shown as circles, while the lines indicate the model biomass scaled by catchability. East Stock and West Stock indices are the spawning stock biomass estimates from operating model 1.

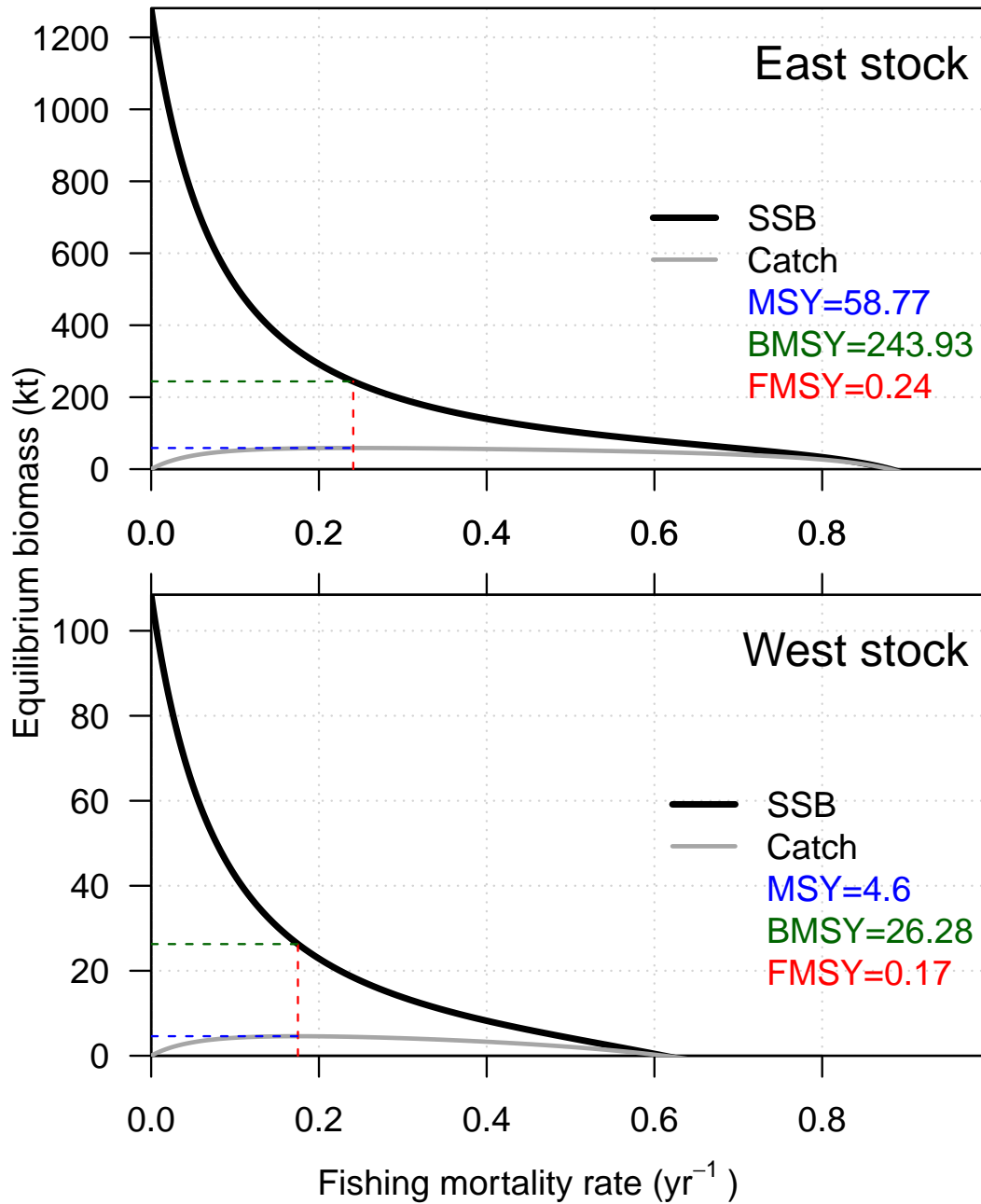


Figure 5: Equilibrium biomass (black) and yield (grey) curves as a function of fishing mortality estimated by the delay difference model fits to the management indices and operating model 1 spawning stock biomass.

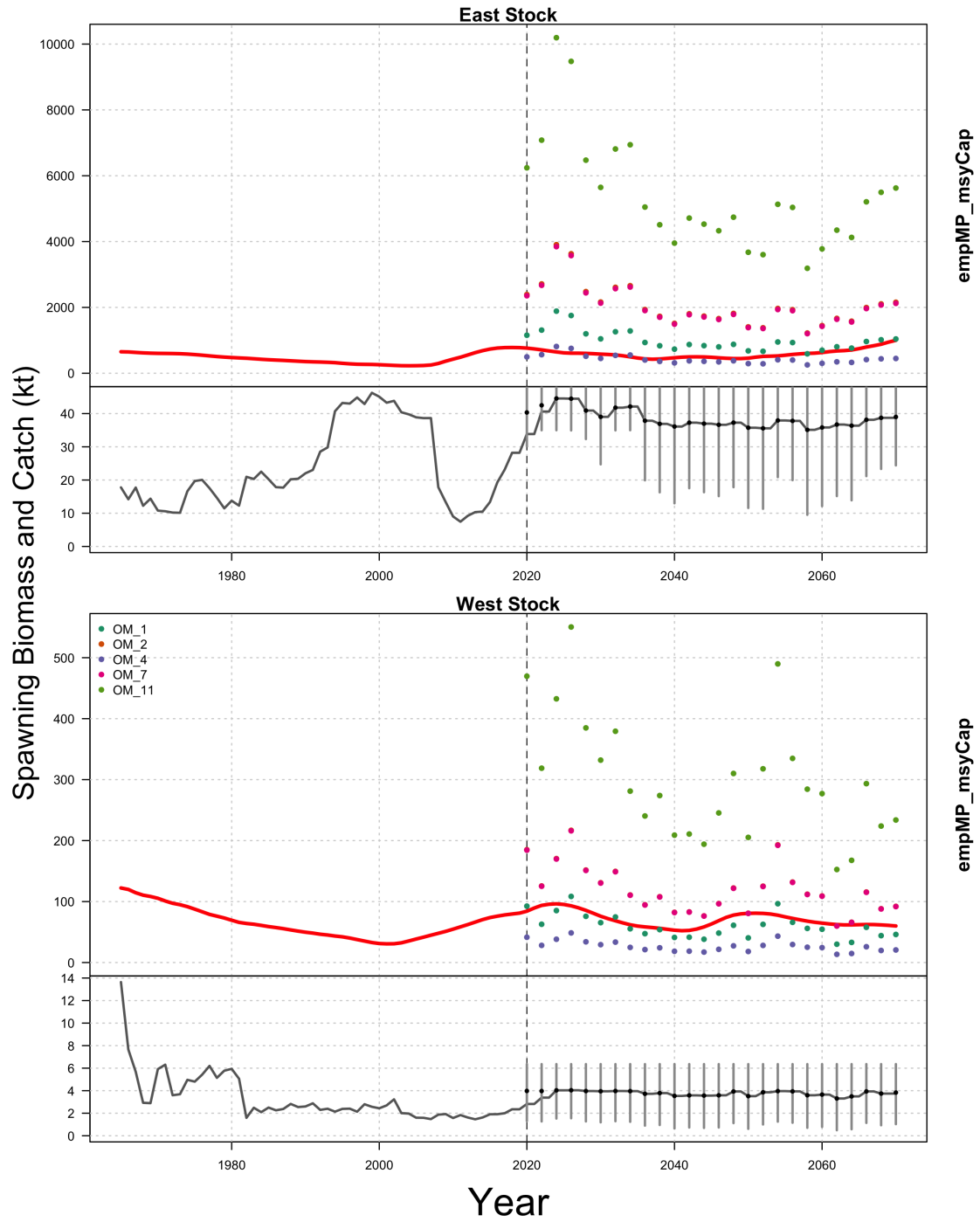


Figure 6: A single simulation replicate from Operating Model 1, showing estimated and simulated spawning stock biomass (coloured dots and red lines), and the area catch (black lines) with mean and central 90% of the distributions of multi-model TAC advice (black points and grey vertical segments in the catch projection period). Each colour of a dot corresponds to the estimate of biomass when using parameters from the tuning subset operating model indicated in the legend.

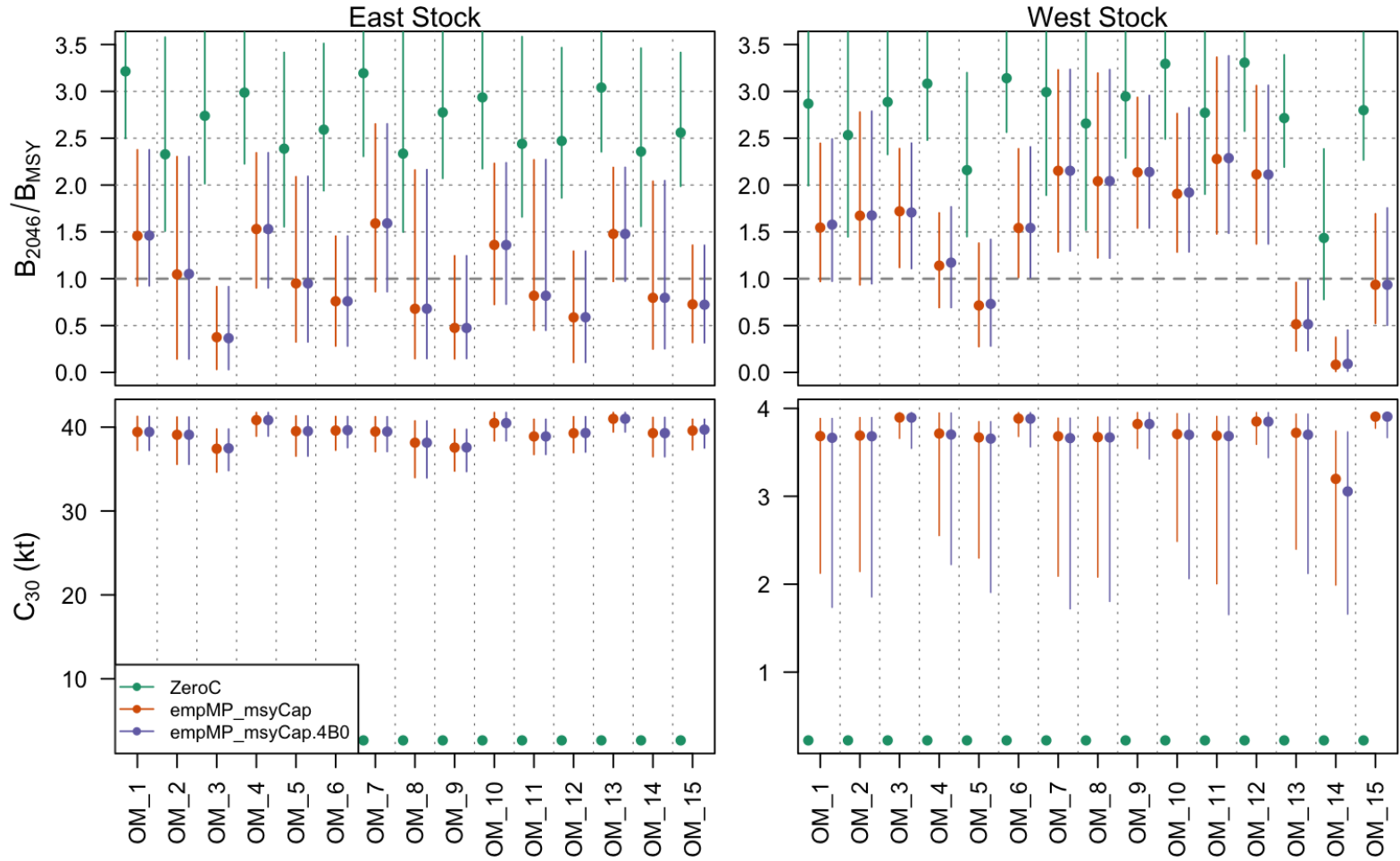


Figure 7: Distributions of MP performance metrics B_{30} and AvC_{30} for empirical MPs with TAC caps set at maximum sustainable yield. Points show the median value of the performance metric, and segments show the central 90% of the distribution over all simulation replicates.

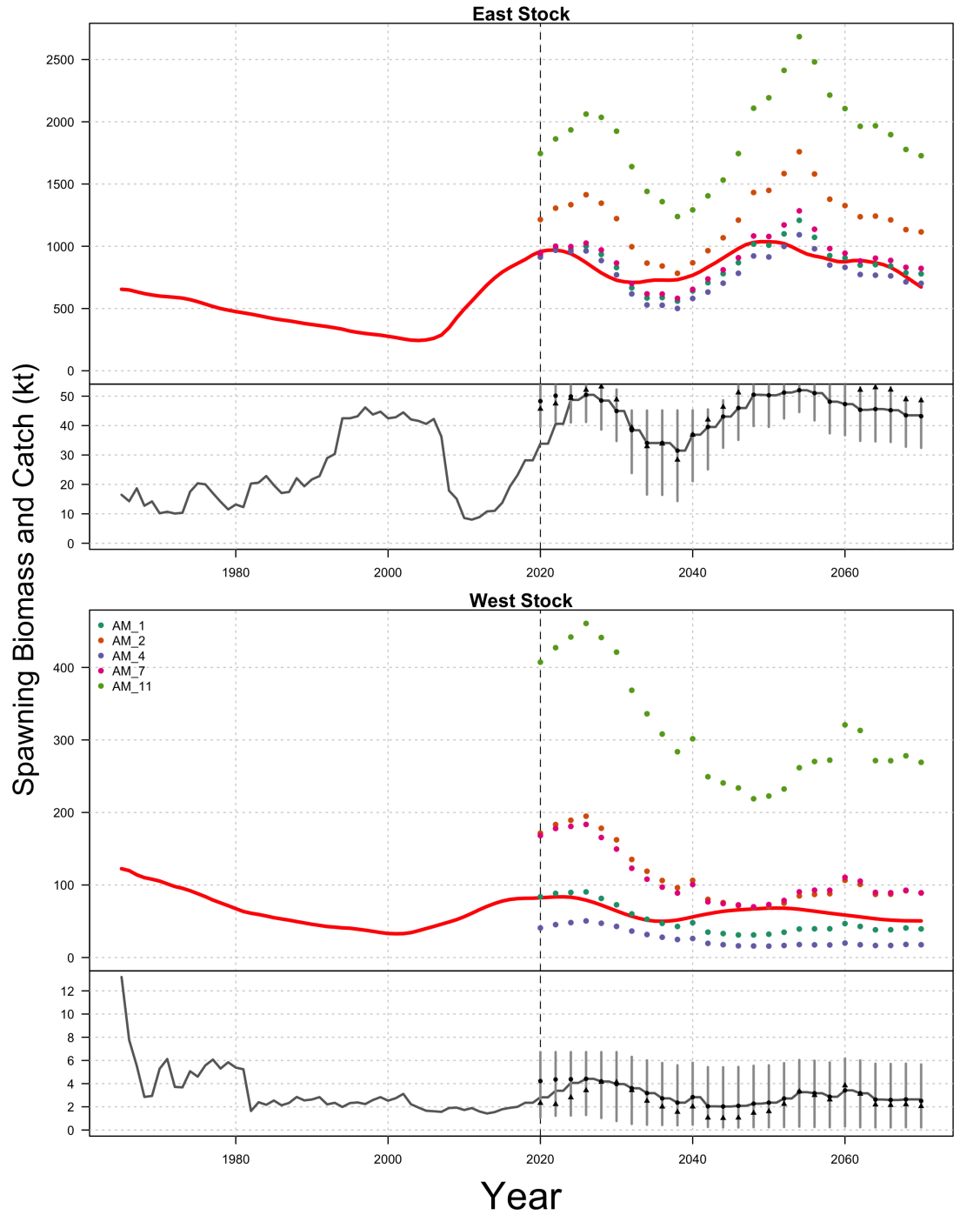


Figure 8: A single simulation replicate from Operating Model 1, showing estimated and simulated spawning stock biomass (coloured dots and red lines), and the area catch (black lines) under the MP_{msyCapF23M.4B0} CMP. In the biomass panels, each colour of a dot corresponds to the estimate of biomass when using AMs tuned to the tuning subset OM indicated in the legend. In the catch panels, multi-model TAC advice is overlaid in the projection period with the mean (circular points), AIC weighted TAC (triangular points), and central 90% of the distributions of sub-procedure TAC (vertical segments) in the projection period.

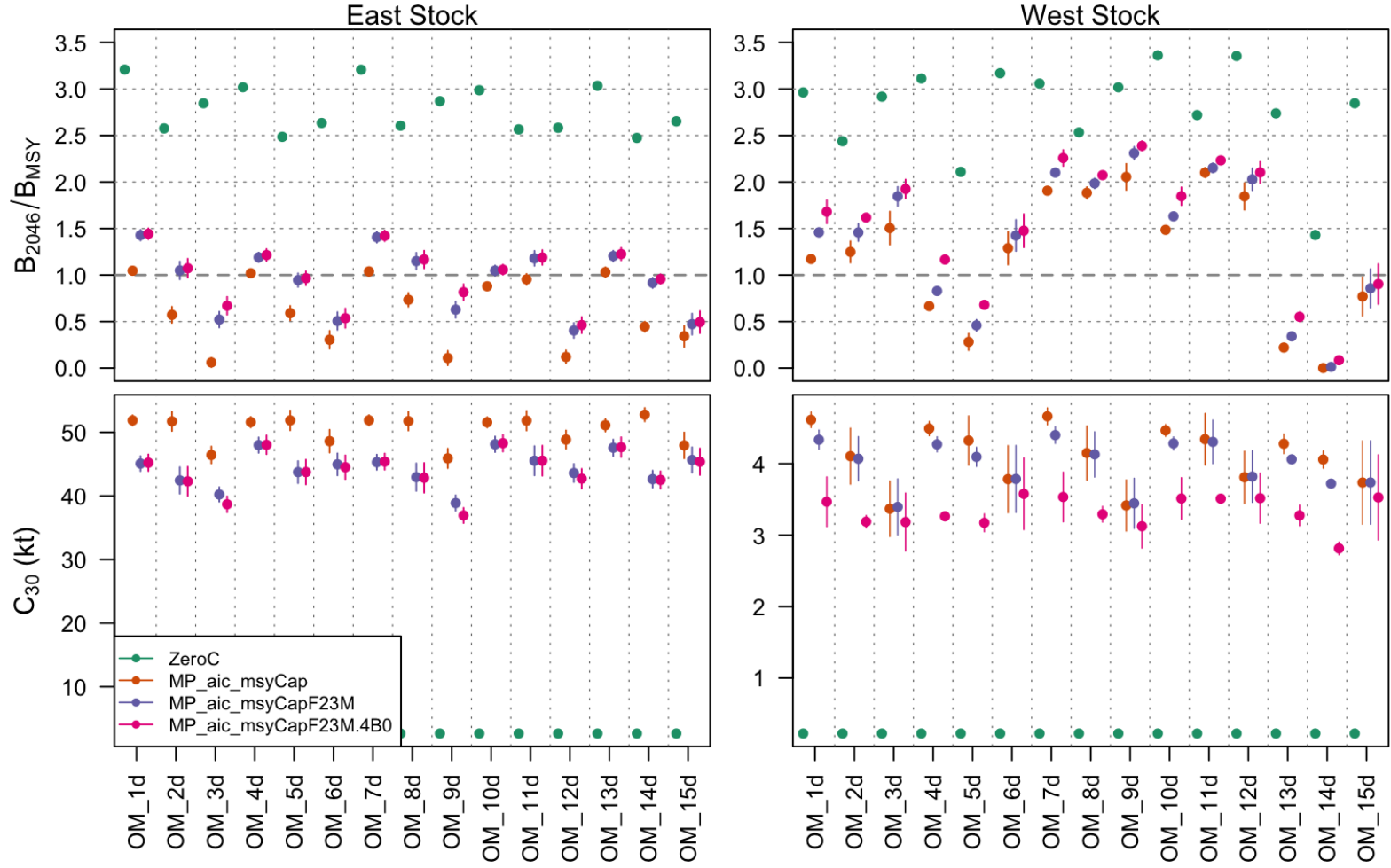


Figure 9: Distributions of MP performance metrics $B_{30} = B_{2046}/B_{MSY}$ and $AvC_{30} = \frac{\sum_{t=2016}^{2046} C_t}{30}$ for Delay Difference MPs with TAC caps set at AM MSY estimates, and TACs weighted by AM AIC values, applied to deterministic operating models. Points show the median value of the performance metric, and segments show the central 90% of the distribution over all simulation replicates.

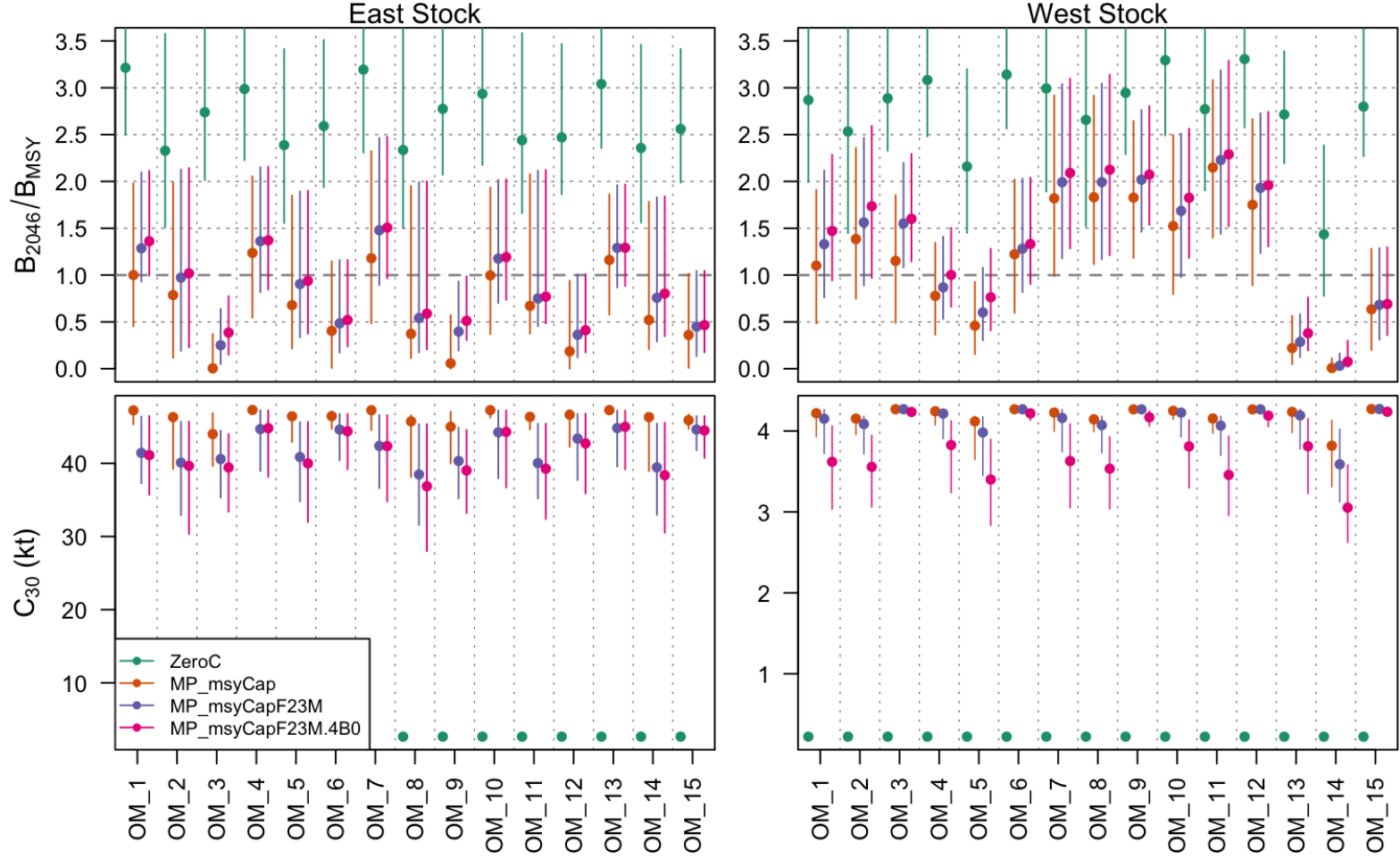


Figure 10: Distributions of MP performance metrics $B_{30} = B_{2046}/B_{MSY}$ and $AvC_{30} = \frac{\sum_{t=2016}^{2046} C_t}{30}$ for Delay Difference MPs with TAC caps set at AM MSY estimates, and TACs weighted equally, applied to operating models 1 through 15. Points show the median value of the performance metric, and segments show the central 90% of the distribution over all simulation replicates.

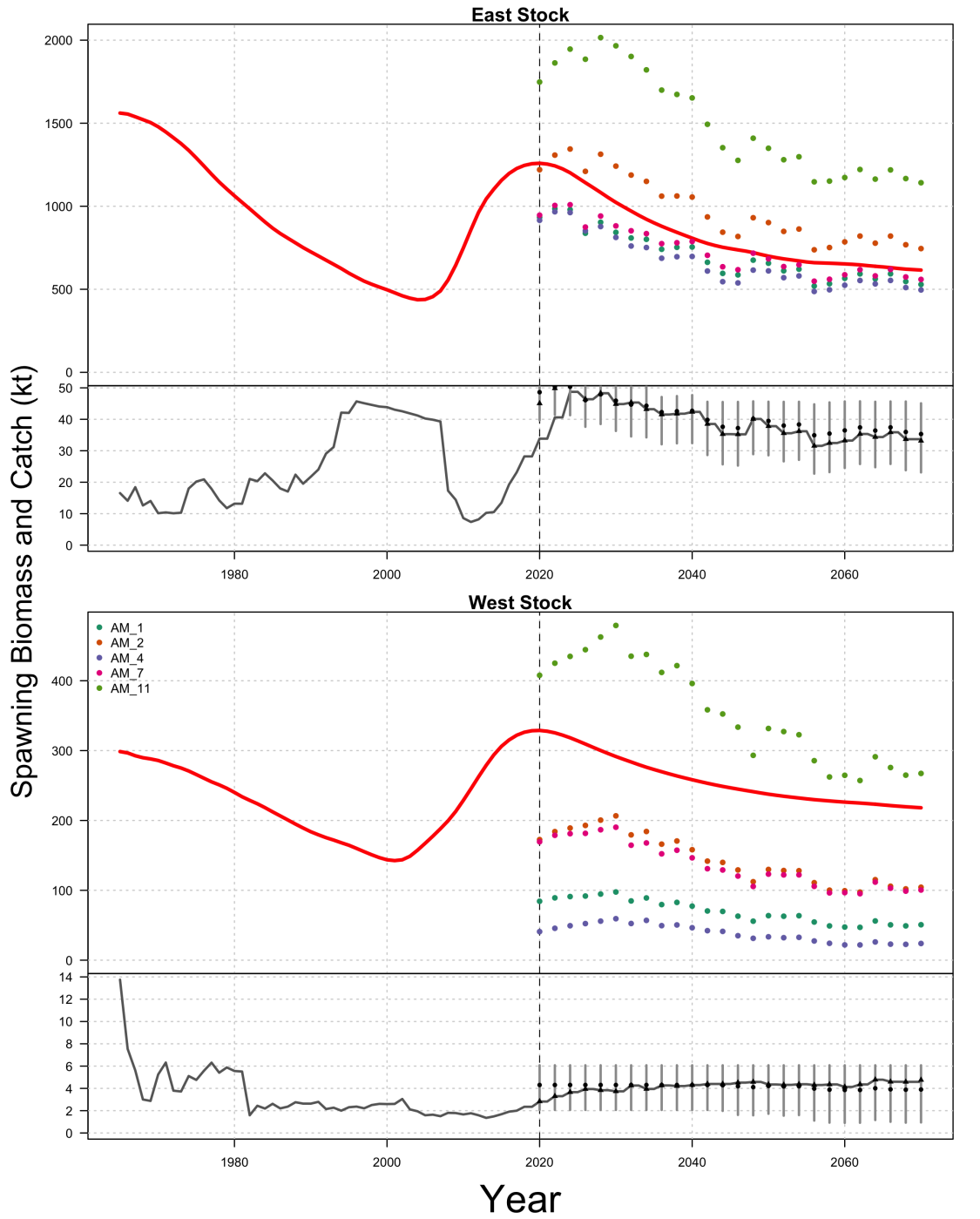


Figure 11: A single simulation replicate from Operating Model 8, which is outside of the tuning subset, showing estimated and simulated spawning stock biomass (coloured dots and red lines), and the area catch (black lines) under the MP_aic_msyCapF23 CMP. In the biomass panels, each colour of a dot corresponds to the estimate of biomass when using AMs tuned to the tuning subset OM indicated in the legend. In the catch panels, multi-model TAC advice is overlaid in the projection period with the mean (circular points), AIC weighted TAC (triangular points), and central 90% of the distributions of sub-procedure TAC (vertical segments) in the projection period.