SCRS/2019/017 Collect. Vol. Sci. Pap. ICCAT, 76(2): 71-99 (2019)

EVALUATION OF AN $F_{0.1}$ MANAGEMENT PROCEDURE USING AN ALTERNATIVE MANAGEMENT STRATEGY EVALUATION FRAMEWORK FOR ATLANTIC BLUEFIN TUNA

M.R. Morse[1], L.A. Kerr[2], S.X. Cadrin[1]

SUMMARY

We demonstrate a management strategy evaluation (MSE) that was designed to complement the ICCAT ABT-MSE tool. Similar to the ABT-MSE tool, ours includes a two-population, spatially-structured operating model that has had input from the ICCAT Atlantic bluefin tuna community over several iterations. Our operating model is conditioned on seasonal movements derived from telemetry as well as ICCAT perceptions of recruitment, fishing mortality, and observation error. Our MSE supports the evaluation of management procedures that involve age-based estimation models, such as the current virtual population analysis and $F_{0.1}$ management procedure adopted by ICCAT. Preliminary results indicate that the $F_{0.1}$ management procedure is sustainable in the medium-term future (20 years), causing an initial decrease in spawning biomass followed by some rebuilding of both western and eastern populations. Relative inter-annual variation in yields was greater for eastern fisheries than western fisheries. This MSE approach will be used along with the ABT-MSE tool to facilitate workshops to gather input from U.S. fishery stakeholders.

RÉSUMÉ

Ce document présente une évaluation de la stratégie de gestion (MSE) conçue pour compléter l'outil ABT-MSE de l'ICCAT. Notre outil est semblable à l'outil ABT-MSE et inclut un modèle opérationnel à deux populations et structuré spatialement qui a bénéficié de la contribution de la communauté de spécialistes du thon rouge de l'Atlantique de l'ICCAT au cours de plusieurs itérations. Notre modèle opérationnel est conditionné par les mouvements saisonniers dérivés de la télémétrie, ainsi que par les perceptions du recrutement, de la mortalité par pêche et des erreurs d'observation de l'ICCAT. Notre MSE soutient l'évaluation des procédures de gestion qui impliquent des modèles d'estimation basés sur l'âge, tels que l'analyse de la population virtuelle actuelle et la procédure de gestion F0,1 adoptées par l'ICCAT. Les résultats préliminaires indiquent que la procédure de gestion F0,1 est durable à moyen terme (20 ans), entraînant une diminution initiale de la biomasse du stock reproducteur, suivie d'un rétablissement partiel des populations de l'ouest et de l'est. La variation interannuelle relative de la production était plus grande pour les pêcheries de l'est que pour les pêcheries de l'ouest. Cette approche MSE sera utilisée avec l'outil ABT-MSE pour faciliter les ateliers afin de recueillir les commentaires des parties prenantes du secteur de la pêche des États-Unis.

RESUMEN

Se muestra una evaluación de estrategias de ordenación (MSE) que se diseñó para complementar la herramienta MSE-ABT de ICCAT. De un modo similar a la herramienta MSE-ABT, nuestra evaluación incluye un modelo operativo espacialmente estructurado de dos poblaciones que ha tenido aportaciones de la comunidad de atún rojo del Atlántico de ICCAT durante varias iteraciones. Nuestro modelo operativo está condicionado por los movimientos estacionales derivados de la telemetría, así como de las percepciones ICCAT de reclutamiento, mortalidad por pesca y error de observación. Nuestra MSE respalda la evaluación de procedimiento de ordenación que incluye modelos de estimación basados en la edad, como el análisis de población virtual actual y el procedimiento de ordenación F0,1 adoptado por ICCAT. Los resultados preliminares indican que el procedimiento de ordenación F0,1 es

[1] University of Massachusetts Dartmouth, School for Marine Science and Technology, 836 Rodney French Blvd, New Bedford, MA, USA, mmorse1@umassd.edu
[2] Gulf of Maine Research Institute, 350 Commercial St, Portland, ME, USA

sostenible en un futuro a medio plazo (20 años), y genera un descenso inicial de la biomasa reproductora seguido de cierta recuperación de las poblaciones oriental y occidental. La variación interanual relativa en los rendimientos fue mayor para las pesquerías orientales que para las occidentales. Este enfoque MSE se utilizará junto con la herramienta MSE ABT para facilitar los talleres para recopilar los datos de entrada de las partes interesadas de la pesquería estadounidense.

KEYWORDS

Atlantic bluefin tuna, management strategy evaluation, simulation, management procedure, virtual population analysis

1. Introduction

Management strategy evaluation (MSE) is the state-of-the-art for testing the performance of management procedures for meeting desired fishery objectives (Hilborn and Walters 1992, Bunnefeld *et al.* 2011). The approach applies closed-loop simulation with an operating model that simulates both the natural and human aspects of the fishery resource system, wherein the simulated status of the resource triggers action based on management strategies, and subsequent management decisions in-turn affect fishing activities and feedback on the resource (Sainsbury *et al.* 2000, Bunnefeld *et al.* 2011). This application of simulation testing can be an effective tool in identifying management strategies that can minimize potential adverse ecological and economic impacts of resource management (Kerr and Goethel 2014).

MSE is particularly useful for testing the impact of uncertainties and assumptions in the stock assessment and management decision-making process on future sustainability of the resource, and for identifying management procedures that are robust to these uncertainties (Carruthers and Hordyk 2018). For example, stock mixing and life history characteristics (e.g., reproductive capacity) are recognized as major sources of uncertainty for informing Atlantic bluefin tuna fisheries management strategies and catch advice (Porch 2005, ICCAT 2015b, ICCAT 2018a). MSE can be used to evaluate the sensitivity of reference points to these uncertainties and to test the performance of associated harvest control rules in achieving fishery management objectives.

ICCAT is currently sponsoring the development of MSE for Atlantic bluefin tuna (ICCAT 2015a). Initial MSE development is being conducted using a multi-stock spatial operating model to test empirical management procedures (Carruthers and Butterworth 2018a,b). As MSE operates at the interface between science and policy, it is important to closely align scientific analysis with management policy decision-making as much as possible (Punt *et al.* 2016). Therefore, the MSE documented here is designed to complement and extend the ICCAT MSE effort by filling analytical gaps utilizing unique operating models to evaluate the performance of the current management procedure used for Atlantic bluefin tuna: virtual population analysis (VPA) with a $F_{0.1}$ harvest strategy.

The goal of this study was to apply simulation testing to evaluate the current ICCAT management procedure in the context of Atlantic bluefin tuna stock mixing. A previously developed operating model for Atlantic bluefin tuna (Kerr *et al.* 2016, 2018) was adapted to run in a medium-term closed-loop simulation to test the VPA and $F_{0.1}$ management procedure on the western and eastern stocks. The operating model and simulation-estimation analyses were revised to address input from the ICCAT Atlantic bluefin tuna assessment groups (Anon. 2013; Kerr *et al.* 2013, 2015; Morse *et al.* 2018b). This application of MSE has been developed in collaboration with the ICCAT community with the aim of advancing the scientific basis of fishery management decisions for this important resource.

2. Methods

2.1 Operating and observation model

An earlier version of the Atlantic bluefin tuna operating and observation model employed in this study is documented in Kerr *et al.* (2016, 2018) and Morse (2018). The operating model was coded and run in R (version 3.5.2, R Core Team 2016), and the model specifications are listed in **Table 1**. The operating model is age structured (ages 1-29) and simulates movement of fish across seven geographic zones, two stock areas (**Figure 1**), and four seasons. The eastern population originates and spawns in the Mediterranean Sea and the western population in the Gulf of Mexico. Bluefin tuna from one stock area move to the other area, but spawn only in their natal area, according to the "overlap" model structure (Porch *et al.* 2001). Fish movement over seven-zones and four seasons was conditioned on fishery-independent telemetry-based movement probability matrices developed by Galuardi *et al.* (2018) using a simulation framework (SatTagSim) that analyzed Atlantic bluefin tuna tagging data. Life history parameters for natural mortality, growth, and spawning fraction were identical to those used in the most recent stock assessment, and the older maturity ogive based on spawning fraction was used for the western population (ICCAT 2018a). All fleet and survey indices of relative abundance used in the most recent stock assessment (ICCAT 2018a) were included in the operating model (**Table 2**).

The operating model was conditioned on ICCAT (2018a) estimates of recruitment and fishing mortality for the period 1974 to 2015 and estimates of abundance-at-age for the year 1974, as described by Morse (2018). The projected period, spanning 21 years from 2016 to 2036, calculated recruitment of age 1 fish to the western and eastern populations using hockey stick stock-recruit models. The hockey stick models (**Figure 2**) were parameterized on ICCAT (2018a) VPA estimates of recruitment and on the "effective" spawning stock biomass (SSB), i.e., the SSB (based on the maturity ogive) in the spawning zones (zone 1, Gulf of Mexico, for the western population, and zone 7, Mediterranean Sea, for the eastern population) during the spawning season (quarter = 1, spring) in the operating model. The justification for using effective SSB is the operating model structure, which allows only mature fish that return to the spawning zones during the spawning season to contribute to reproduction. The stock-recruitment relationship was time-invariant. Hockey stick parameters R_{max} and SSB* were calculated empirically based on the ICCAT (2015b) methods, where R_{max} is the average R over the entire time series, and SSB* is the average SSB over the six-year range when SSB was the lowest. Annual recruitment was produced with lognormal error structure.

Operating model outputs for the bluefin tuna resource size were represented both from a *stock view*, referring to the geographically-distinct western and eastern mixed-population stocks separated by the 45°W meridian as defined by ICCAT (1981), and a *population view*, referring to the genetically-distinct western and eastern populations originating in their respective natal grounds (Kerr *et al.* 2018; note that the *stock view* is referred to as *area*, and *population view* is referred to as *stock* in the ABT-MSE tool). Stock view attributes were derived by summing the abundance or biomass of fish over all geographic zones contained in a stock area (zones 1 to 3 for the western stock area and 4 to 7 for the eastern stock area; e.g., eq. 1 for stock abundance), and population view attributes were derived by summing over all fish that originated in the respective spawning areas (Gulf of Mexico for the western population and Mediterranean Sea for the eastern population; e.g., eq. 2 for population abundance).

(1) $$\sum_{\substack{West\ z=1:3 \\ East\ z=4:7}} N_{y,a,z,q,p}$$

(2) $$\sum_{\substack{West\ p=2 \\ East\ p=1}} N_{y,a,z,q,p}$$

Fishery pseudodata for catch-at-age, indices of relative abundance, and relative age composition of indices were generated with lognormal observation error (as documented in Kerr *et al.* 2018 and Morse 2018). Whereas 27 abundance indices were included in the historical period (1974-2015), only 11 were continued into the projected period (6 in the west and 5 in the east, **Table 2**) because they had data in the last year of the historical period. The observation model generated stochastic pseudodata with random normal observation error (ε; see derivation below), which included catch-at-age ($C_{y,a,s}$, eq. 3), indices of relative abundance ($I_{y,g}$, eq. 4), and relative age composition of indices ($X_{y,a,g,s}$, eq. 5):

(3) $$C_{y,a,s} = \left(\sum_{\substack{West\ z=1:3 \\ East\ z=4:7}} \sum_{q=1}^{4} \sum_{p=1}^{2} N_{y,a,z,q,p} \frac{F_{y,a,z,q}}{F_{y,a,z,q} + M_{a,q,p}} \left[1 - e^{-(F_{y,a,z,q}+M_{a,q,p})}\right] \right) e^{\varepsilon_{y,a,s}}$$

(4) $$I_{y,g} = \left(\sum\nolimits_{\substack{West\ z=1:3 \\ East\ z=4:7}} \sum_{a=1}^{29} \sum_{p=1}^{2} S_{a,g} N_{y,a,z,q,p} W_{a,p} Q_g \right) e^{\varepsilon_{y,g}}$$

(5) $$X_{y,a,g,s} = \left(\sum\nolimits_{\substack{West\ z=1:3 \\ East\ z=4:7}} \sum_{q=1}^{4} \sum_{p=1}^{2} N_{y,a,z,q,p} \frac{E_{y,g} Q_g S_{a,g}}{E_{y,g} Q_g S_{a,g} + M_{a,q,p}} \left[1 - e^{-(E_{y,g} Q_g S_{a,g} + M_{a,q,p})}\right] \right) e^{\varepsilon_{y,a,s}}$$

where N is abundance, F is fishing mortality rate, M is natural mortality rate, S is selectivity, W is weight, Q is catchability, and E is effort, which are disaggregated by year y, age a, geographic zone z, fleet g, seasonal quarter q, population p, and stock s (for equation derivations see Kerr *et al.* 2018). The operating model produced age-structured catch data for ages 1 to 29, but because the western and eastern Atlantic bluefin tuna stock assessment models have age 16+ and age 10+ groups, respectively, the catch-at-age data were summed across all ages from the plus groups to age 29.

Fishery-dependent and -independent indices of relative abundance were derived from simulated fishing fleets and surveys that emulated the geographic scope, magnitude, and time frame of index data used in the 2017 ICCAT stock assessments. Abundance values used in the calculation of indices (eq. 4) were assumed to be from the beginning of the third quarter (to reflect the fall season when the majority of fishing effort occurs in the Atlantic bluefin tuna fishery), except for indices that measured relative abundance in spawning areas (zones 1 and 7), in which case the abundance was assumed to be from the beginning of the first quarter (to reflect the spawning season). Pseudodata for the relative age composition of indices were derived from the relative fishing mortality rate-at-age within fleets and years, where the index-specific fishing mortality rate $F_{y,a,g}$ was derived as the product of the index-specific effort $E_{y,g}$, catchability Q_g, and selectivity $S_{a,g}$, which were based on 2017 ICCAT VPA estimates (eq. 5). These age composition pseudodata are used by the estimation model to inform the age-selectivities of the corresponding indices of relative abundance (Porch 2003).

Assuming that catch, index, and relative age composition data were lognormally-distributed, observation error (ε) was generated stochastically according to the normal distribution, $\varepsilon \sim N(0, \sigma^2)$. The observation error $\varepsilon_{y,a,s}$ for the catch-at-age of each stock s and the relative age composition of each index had a standard deviation σ_s calculated as the root mean square error of the observed and predicted catch values x_i and $\hat{x}_i$ over all ages a to A and years y to Y:

(6) $$\sigma_s = \sqrt{\frac{1}{Y}\frac{1}{A}\sum\nolimits_{y}^{Y}\sum\nolimits_{a}^{A}(\ln x_i - \ln \hat{x}_i)^2}$$

The observed and predicted catch-at-age values were derived from an exploratory age-structured assessment program (ASAP, Legault and Restrepo 1998) analysis of Atlantic bluefin tuna data (Maguire *et al.* 2018). The observation error $\varepsilon_{y,g}$ for each index of relative abundance g had a standard deviation σ_g calculated as the root mean square error over all years of the index time series:

(7) $$\sigma_g = \sqrt{\frac{1}{Y}\sum\nolimits_{y}^{Y}(\ln x_i - \ln \hat{x}_i)^2}$$

The observed and predicted index values x_i and $\hat{x}_i$ were derived from the western and eastern bluefin tuna VPAs (ICCAT 2018a).

2.2 Management procedure

The simulated management procedure consisted of application of estimation models, estimation of reference points, and implementation of target fishing mortality. Corresponding to the average ICCAT assessment-management cycle timing for bluefin tuna, the assessment-management cycle repeated every three years for seven projected cycles, resulting in a total of 21 projected years. Each assessment used only the past 42 years of available fishery data. One hundred realizations of the MSE were run. Runs in which one or more estimation models did not converge were excluded from the results.

2.2.1 Stock assessment model

The estimation models were the single stock VPAs used for past Atlantic bluefin tuna stock assessments (VPA-2BOX version 4.01, Porch *et al.* 2001). The estimation model was applied separately to pseudodata for each stock. For the historical period, the estimation model had same settings as Morse (2018). However, the following model settings were changed for the projected periods to increase convergence rates during automated application:

- 8-year recruitment penalty
- No selectivity penalty
- F-ratio parameters estimated in 5-year blocks of constrained random walk (east stock only)

These settings were based on Zarrad *et al.*'s (2018) review of the 2014 eastern bluefin tuna stock assessment. In addition, variance scaling parameters (index weighting) were fixed at the true values for the estimation models in both the historic and projected periods in order to increase convergence rates to achieve the desired 21-year projections and assessment cycles.

2.2.2 $F_{0.1}$ reference point

The perceived $F_{0.1}$ for each stock and assessment-management cycle was calculated from the estimation model results using the yield-per-recruit (*ypr*) function in the fishmethods R package (Nelson 2017). The input partial recruitment to the fishery $P_{a,s}{}'$ was derived from the fishing mortality rates-at-age for the last three years of the VPA results $F_{a,y,s}$ excluding the final year (eq. 8-11). The true weight-at-age and maturity-at-age were assumed known without error. Because this value of $F_{0.1}$ was calculated based on partial recruitment scaled to apical F (F_{full}), it was then scaled by the average partial recruitment of the fully-recruited ages, $\bar{P}_{a_{ref},p}{}'$, termed "reference ages" (ages where partial recruitment $P_{a,y,p}$ was greater than or equal to 0.8 in 2012-2014), to produce the desired $F_{0.1}$ (eq. 12).

(8) $$F_{full_{y,s}} = \max_a(F_{a,y,s})$$

(9) $$P_{a,y,s} = {F_{a,y,s}} \Big/ {F_{full_{y,s}}}$$

(10) $$P_{a,s} = \frac{\sum_y^{y+2} P_{a,y,s}}{(y+2) - y + 1}$$

(11) $$P_{a,s}{}' = \frac{P_{a,s}}{\max_a(P_{a,s})}$$

(12) $$F_{0.1} = F_{0.1_{apical}} * \bar{P}_{a_{ref},p}{}'$$

For each three-year assessment-management cycle, the target $F_{0.1}$ for each stock $F_{0.1,s}$ was scaled by the selectivity-at-age (based on the last year of the historical period) for each geographic zone of the operating model $S_{a,z}$ (eq. 13). The resulting fishery exploitation rates $F_{a,z,s}$ were applied uniformly to the operating model for all three years of the assessment-management cycle, assuming no changes in harvest levels between assessments.

(13) $$F_{a,z,s} = F_{0.1,s} S_{\substack{a,west\ z=1:3 \\ east\ z=4:7}}$$

The MSE assumed perfect implementation of the $F_{0.1}$ management procedure (i.e., no implementation error), such that the value of $F_{0.1}$ estimated by the VPA was applied directly to the simulated fishery. The estimated values of $F_{0.1}$ were compared to the true (operating model) values of $F_{0.1}$ for the population and stock views to assess whether there was any estimation error.

The true values of $F_{0.1}$ were calculated in the same manner as the perceived values of $F_{0.1}$, but the fishing mortality rates used to calculate the partial recruitment differed. For the population view, the abundance-at-age by year for each population p in the first quarter was calculated by summing over all seven geographic zones:

(14) $$N_{a,y,q=1,p} = \sum_{z=1}^{7} N_{a,y,z,q=1,p}$$

The annual instantaneous fishing mortality rate $F_{a,y,p}$ for each population was calculated using the abundance $N_{a,y,p}$ and natural mortality rate $M_{a,p}$:

(15) $$F_{a,y,p} = \ln\left(\frac{N_{a,y,p}}{N_{a+1,y+1,p}}\right) - M_{a,p}$$

For the stock view, the operating model fishing mortality rates $F_{a,y,z}$ were summed over geographic zones to get the fishing mortality rates for the western stock (zones 1-3) and eastern stock (zones 4-7), respectively:

(16) $$F_{a,y,s} = \sum\nolimits_{\substack{West\ z=1:3 \\ East\ z=4:7}} F_{a,y,z}$$

For both the population and stock views, the average partial recruitment $P_{a,p}'$ over the last three years of the current projected period excluding the final year were calculated (eq. 8-11) for input to the yield-per-recruit function, then scaled to the average partial recruitment of the respective reference ages (eq. 12).

2.3 Performance metrics

MSE performance was evaluated based on time series of population SSB, stock SSB, and yield, on the frequency of SSB falling below the SSB* hockey stick hinge point, and on estimation error in the $F_{0.1}$ reference point. Additional performance metrics were chosen to emulate those used in the ABT-MSE tool (Carruthers 2019) to enhance comparability. They are listed and defined in **Table 3.**

3. Results

3.1 Long-term deterministic outcomes

Deterministic long-term equilibrium SSB and yield curves generated for a range of F values were characterized by sharp peaks and wavelike patterns (**Figure 3**). The peaks in the SSB curves indicate the hinge point in the hockey stick model (**Figure 2**), where the number of recruits per spawner decrease at a constant rate.

The peak in the yield curves is the maximum sustainable yield (MSY), which occurs at approximately F = 0.14 for the western population and F = 0.29 for the eastern population. The sharp peak and abrupt drop are characteristic of yield curves for stocks modeled using a hockey stick stock-recruit relationship (Mesnil and Rochet 2010). The abrupt drop is F_{crash}, where yield drops off drastically. In this case, $F_{MSY} = F_{crash}$, so effort limits set at F_{MSY} are highly risky and unsustainable.

A second peak was evident in the yield curve for the western stock at F = 0.29, which is the F_{crash} for the eastern population, and shows up because the western stock is composed of both the western population and migrants from the eastern population. Therefore, although yields of the western population above F = 0.14 are unsustainable, relatively high yields of the western stock may sustained up until F = 0.29 because of the influx of eastern fish but run the risk of crashing the western population.

3.2 Stock assessment model performance

Previous research demonstrated through simulation testing of the Atlantic bluefin tuna VPA that model estimates of spawning stock biomass are affected by mixing among the western and eastern populations across the western and eastern stock management areas (Morse 2018). This research showed that the VPA significantly overestimated western recruitment (~200% positive bias) but underestimated eastern recruitment (~30% negative bias). Similarly, spawning stock biomass was underestimated for the eastern population (~70% negative bias) but overestimated for the western population (~100% positive bias).

The convergence rate of the stochastic MSE for 21-year projections was 38% (defined as the percentage of runs in which all VPA models converged up until the last year of the projection—a MSE realization aborted if either the eastern or western VPA models did not converge). Lack of VPA convergence generally resulted from terminal F parameters hitting boundary constraints, which most often occurred in the historical period or the first projected assessment period.

Stock selectivities were relatively well estimated on average, but there was a large variability in the estimated values (**Figures 4** & **5**). Of the indices that were carried into the projected period, selectivities for most were well estimated, except the Canadian Gulf of St Lawrence acoustic survey (CAN_GSL_Acoustic), the western Japanese longline (JLL_RECENT), and the western Mediterranean larval survey (WMED_LARV; **Figure 6**). The relative age composition of catch for the end of the historic period (**Figure 7**) and end of the projected period (**Figure 8**) indicated the relative contribution of fleets to catch volume. The high volume of catch from the western Japanese longline fleet at the end of the projected period (**Figure 8**), in particular, paired with the large spread and negatively biased medians of selectivity-at-age estimates (**Figure 6**), suggest that poor estimation of this fleet contributes to challenges in VPA estimation of fishing mortality rates.

3.3 $F_{0.1}$ management procedure performance

Fishing at the $F_{0.1}$ level initially caused a decline in the average SSB of the western and eastern populations (and, by extension, the western and eastern stocks) followed by a partial rebuilding to levels among the highest in the recent historical period (**Figure 9**). The initial decline in projected SSB may be a relic of the weak eastern recruitments in the late 2000s and weak western recruitments in the early 2010s (ICCAT 2018a). SSB rebuilds when the small cohorts begin to be replaced by stronger ones as the recruits modeled by the hockey stick model being to reach maturity, but rebuilding may be sensitive to the parameterization of the hockey stick model. As the model assumes that western fish reach maturity later than eastern fish, the eastern population SSB rebuilds faster and the western lag time is longer. The western stock SSB is over three times greater than population SSB because of the influx of eastern fish.

Average projected western and eastern stock yields increased from the most recent years of the historical period (up until 2015) under the $F_{0.1}$ management strategy (**Figure 10**). Particularly high yields in the first ~5 projected years explain the large initial drops in SSB (**Figure 9**). Average yields leveled off in the final year of projections to approximately 6000 tonnes for the western stock and 25,000 tonnes for the eastern stock. Median inter-annual variation in yield was greater for the eastern stock (45% relative difference) than the western stock (15% relative difference) (**Figure 11**).

Relative to the deterministic equilibrium, SSB_0, western SSB was never depleted (SSB/SSB_0 always > 1) but eastern SSB was always depleted after both the first and second 10-year projected management procedure application periods (**Figure 12**). Relative to the zero catch trajectory, SSB of both stocks was always depleted at the end of the projected management procedure application period (**Figure 13**). Effective SSB of the western population was below the hinge point in the final projection year in 23% of realizations, implying diminishing recruits according to the hockey stick model (**Figure 14**). Effective SSB was always above the hinge point for the eastern population.

The MSE suggests that estimation of $F_{0.1}$ was reliable, given that median estimation bias in $F_{0.1}$ was low, except for overestimation in the historical period (**Table 4, Figure 15**).

4. Discussion

In general, this study demonstrated that fishing at $F_{0.1}$ could be sustained for Atlantic bluefin tuna in the medium-term future, with high inter-annual variation in eastern fishery yields. Relative to recent ICCAT TACs for 2018 to 2020, medium-term fishery yields were higher in the western fishery and lower in the eastern fishery (ICCAT SCRS 2017).

Although estimation of $F_{0.1}$ reference points was generally accurate, this study also raised general concerns about estimating reference points for mixed stocks. The conventional method of using separate western and eastern stock assessments to estimate respective $F_{0.1}$ reference points has the potential to produce inaccurate values because true natural mortality, growth, and partial recruitment differ amongst the populations making up the mixed stocks, whereas management typically assumes these parameters are constant among all fish in the given stock. Previous studies have demonstrated that in the presence of spatial structure, such as spatial variation in fishing pressure

(Hart 2001), metapopulation structure (Ying *et al.* 2011) or a range of possible spatial population structures (Goethel and Berger 2017), estimation of yield-per-recruit and associated reference points ($F_{0.1}$, F_{MSY}) may misrepresent the true values.

Evaluating the performance of model-based management strategies, as opposed to empirical strategies (e.g., survey trends), is a challenging task. Testing the VPA-2BOX estimation model and $F_{0.1}$ management procedure through MSE required an age-structured operating model and several assumptions about model specifications and parameterization. It was not possible to implement management procedures based on a VPA or other age-based assessment methodology in the current ICCAT MSE iteration, because simulating age-structured or length-structured catch data is a complex issue. Some approximations were considered (e.g., Rice 2018), but the subgroup concluded that mimicking the VPA-$F_{0.1}$ management procedure implemented in 2017 was not appropriate at this stage due to the lack of generated age-structured pseudodata with error needed to test such models (ICCAT 2018b). To simplify the MSE process, which is generally complicated for statistical stock assessment models (Punt 1997), the operating model in this study was not based on fitting to the available data but was conditioned on parameters estimated externally by the VPA-2BOX model in the 2017 bluefin stock assessment (ICCAT 2018a).

Previous studies have avoided simulation testing management procedures based on statistical VPA because of convergence problems due to assumptions made about specifications (such as index weighting), which is usually an extensive iterative process involving visual examination of results for alternative sets of specifications and making educated choices (Punt 1997). This iterative process must be automated in MSE, and thus fixed parameters and the starting values of estimated parameters must be robust to a range of potential pseudodata values and error structures. Even so, simulating a management strategy involving a statistical stock assessment model "runs the risk that fully automated fitting procedures may not find the global minimum that would be detected in the comprehensive searches typical of 'best assessment' approaches" (Punt *et al.* 2016). By necessity, the approach in this study over-simplified the actual stock assessment process. However, MSE exercises that do not simulate the actual assessment method, although allowing a broader set of management strategies to be explored more quickly, "risk that the actual error distribution... does not match that assumed, and hence the values of the performance statistics are incorrect" (Punt *et al.* 2016).

One limitation of this simulation approach was the assumption of stationarity in population, fishery, and environmental factors in projections. For example, fish movement parameters were time-invariant both in the historical and projected periods, which is an oversimplification of population dynamics as evidenced by otolith chemistry data (Morse *et al.* 2018a). Failletaz *et al.* (2019) showed that large-scale shifts in climatic conditions may drive changes in the regional distribution of bluefin tuna in the Atlantic and influencing perceptions of stock sizes and appropriate quotas. The operating model documented here assumed the older spawning fraction for the western stock, but a competing theory of younger age at spawning with younger fish spawning the Slope Sea (Richardson *et al.* 2016) presents compelling evidence for testing additional operating models that evaluate management strategy performance under this alternative scenario. Future applications of this MSE approach will explore some of these additional scenarios.

To inform further application of this MSE approach, a series of fishery stakeholder engagement workshops are scheduled for later this year. U.S. stakeholders will be briefed on the MSE process and will be prompted for input on fishery management objectives, operating model scenarios, performance indicators, and candidate management procedures. This input will be used to expand the scope of analyses using the ICCAT ABT-MSE tool and the tool documented here.

Acknowledgements

This work was funded by the US NOAA Bluefin Tuna Research Program and benefitted from collaboration with Ben Galuardi, Clay Porch, John Walter, Matt Lauretta, Nathan Taylor, Tristan Rouyer, Ai Kimoto, J.-J. Maguire, Dave Secor, and Doug Butterworth.

References

Ailloud, L.E., Lauretta, M.V., Hanke, A.R., Golet, W.J., Allman, R.J., Siskey, M.R., Secor, D.H., and Hoenig, J.M. 2017. Improving growth estimates for Western Atlantic bluefin tuna using an integrated modeling approach. Fish. Res. 191: 17-24.

Anonymous. 2013. Report of the 2013 Meeting on Bluefin Stock Assessment Methods (Gloucester, Massachusetts, United States – July 20 to 22, 2013)

Bunnefeld, N., Hoshino, E., and Milner-Gulland, E.J. 2011. Management strategy evaluation: a powerful tool for conservation? Trends Ecol. Evolut. 26: 9.

Carruthers, T. 2019. ABTMSE: Atlantic Bluefin Tuna Management Strategy Evaluation. R package version 4.1.1.

Carruthers, T. and Butterworth, D. 2018a. Summary of a reference set of conditioned operating models for Atlantic bluefin tuna. ICCAT Collect. Vol. Sci. Pap. 74(6): 3523-3541.

Carruthers, T. and Butterworth, D. 2018b. Performance of example management procedures for Atlantic bluefin tuna. ICCAT Collect. Vol. Sci. Pap. 74(6): 3542-3552.

Carruthers, T.R., and Hordyk, A.R. 2018. The Data-Limited Methods Toolkit (DLMtool): An R package for informing management of data-limited populations. Methods Ecol. Evolut. doi:10.1111/2014-210X.13081.

Cort, J.L. 1991. Age and growth of the bluefin tuna, *Thunnus thynnus* (L.) of the northeast Atlantic. ICCAT Collect. Vol. Sci. Pap. 35(2): 213-230.

Faillettaz, R., Beaugrand, G., Goberville, E., and Kirby, R.R. 2019. Atlantic Multidecadal Oscillations drive the basin scale distribution of Atlantic bluefin tuna. Sci. Adv. 5(1): eaar6993. doi:10.1126/sciadv.aar6993.

Galuardi, B., Cadrin, S.X., Arregi, I., Arrizabalaga, H., DiNatale, A., Brown, C., Lauretta, M., and Lutcavage, M. 2018. Atlantic bluefin tuna area transition matrices estimated from electronic tagging and SATTAGSIM. ICCAT Collect. Vol. Sci. Pap. 74(6): 2903-2921.

Goethel, D.R., and Berger, A.M. 2017. Accounting for spatial complexities in the calculation of biological reference points: effects of misdiagnosing population structure for stock status indicators. Can. J. Fish. Aquat. Sci. 74: 1878-1894.

Hart, D.R. 2001. Individual-based yield-per-recruit analysis, with an application to the Atlantic sea scallop, *Placopecten magellanicus*. Can. J. Fish. Aquat. Sci. 58: 2351-2358.

Hilborn, R., and Walters, C.J. 1992. Quantitative Fisheries Stock Assessment: Choice, Dynamics and Uncertainty: Springer, Medford, MA.

ICCAT. 1981. Recommendation on bluefin management measures 81-01. ICCAT, Madrid, Spain. Available from https://www.iccat.int/Documents/Recs/compendiopdf-e/1981-01-e.pdf [accessed 1 November 2017].

ICCAT. 1999. Report of the ICCAT SCRS bluefin tuna stock assessment session. ICCAT Collect. Vol. Sci. Pap. 49(2): 1-191.

ICCAT. 2015a. Recommendation by ICCAT on the development of harvest control rules and of management strategy evaluation. 15-07.

ICCAT. 2015b. Report of the 2014 Atlantic bluefin tuna stock assessment session. ICCAT Collect. Vol. Sci. Pap. 71(2): 692-945.

ICCAT. 2018a. Report of the 2017 ICCAT bluefin stock assessment meeting. ICCAT Collect. Vol. Sci. Pap. 74(6): 2372-2535.

ICCAT. 2018b. Report of the 2018 ICCAT Bluefin Tuna Species Group MSE Intersessional Meeting (Madrid, Spain 16-20 April 2018).

ICCAT Standing Committee on Research and Statistics (SCRS). 2017. Report of the Standing Committee on Research and Statistics (SCRS). ICCAT, Madrid, Spain.

Kerr, L.A., and Goethel, D.R. 2014. Simulation modeling as a tool for synthesis of stock identification information. *In* Stock Identification Methods: Applications in Fishery Science, Second Edition. *Edited by* S.X. Cadrin, L.A. Kerr, and S. Mariani. Elsevier, Waltham, MA, USA pp. 502-533.

Kerr L.A., Cadrin, S.X., Secor, D.H. and Taylor, N. 2013. A simulation tool to evaluate effects of mixing between Atlantic bluefin tuna stocks. ICCAT Collect. Vol. Sci. Pap. 69(2): 742-759.

Kerr LA, Cadrin, S.X., Secor, D.H. and Taylor, N. 2015. Evaluating the effect of Atlantic bluefin tuna movement on the perception of stock units. ICCAT Collect. Vol. Sci. Pap. 71(4): 1660-1682.

Kerr, L.A., Cadrin, S.X., Secor, D.H., and Taylor, N. 2016. Modeling the implications of stock mixing and life history uncertainty of Atlantic bluefin tuna. Can. J. Fish. Aquat. Sci. 74(11): 1990-2004.

Kerr, L.A., Morse, M.R., Cadrin, S.X., and Galuardi, B. 2018. Application of an Atlantic bluefin tuna operating model to generate pseudodata for stock assessment testing. ICCAT Collect. Vol. Sci. Pap. 74(6): 3405-3426.

Legault, C.M., and Restrepo, V.R. 1998. A flexible forward age-structured assessment program. ICCAT Collect. Vol. Sci. Pap. 49(2): 246-253.

Maguire J.-J., Cadrin, S.X., Hanke, A., and Melvin, G. 2018. An exploration of bluefin tuna data in the North Atlantic (West and East + Mediterranean) with ASAP. ICCAT Collect. Vol. Sci. Pap. 74(6): 3205-3223.

Mesnil, B. and Rochet, M-J. 2010. A continuous hockey stick stock-recruit model for estimating MSY reference points. ICES J. Mar. Sci. 67: 1780-1784.

Morse, M.R. 2018. Misspecifying the Maguro: Evaluating stock assessment performance on mixed stocks of Atlantic bluefin tuna. M.Sc. thesis, Department of Fisheries Oceanography, University of Massachusetts Dartmouth, Dartmouth, MA. Available at ResearchGate.

Morse, M., Cadrin, S., Kerr, L., Secor, D., Siskey, M., Arrizabalaga, H., Hanke, A., and Porch, C. 2018a. An updated analysis of bluefin tuna stock mixing. ICCAT Collect. Vol. Sci. Pap. 74(6): 3486-3509.

Morse, M.R., Kerr, L.A., and Cadrin, S.X. 2018b. Simulating virtual population analysis of Atlantic bluefin tuna stocks. ICCAT Collect. Vol. Sci. Pap. 74(6): 3427-3441.

Nelson, G.A. 2017. fishmethods: Fishery Science Methods and Models in R. R package version 1.10-4. Available from https://CRAN.R-project.org/package=fishmethods. [accessed 1 August 2018].

Porch, C.E. 2003. VPA-2BOX Version 3.01 User’s Guide. NOAA Sustainable Fisheries Division Contribution SFD/2003-2004, Miami, FL.

Porch, C.E. 2005. The sustainability of western Atlantic bluefin tuna: a warm-blooded fish in a hot-blooded fishery. Bull. Mar. Sci. 76(2): 363-384.

Porch, C.E., and Hanke, A. 2018. Estimating the fraction of western Atlantic bluefin tuna that spawn by age from size frequency data collected on the Gulf of Mexico spawning grounds. ICCAT Collect. Vol. Sci. Pap. 74(6): 3224-3233.

Porch, C.E., Turner, S.C., and Powers, J.E. 2001. Virtual population analysis of Atlantic bluefin tuna with alternative models of transatlantic migration: 1970-1997. ICCAT Collect. Vol. Sci. Pap. 52(3): 1022-1045.

Punt, A.E. 1997. The performance of VPA-based management. Fish. Res. 29: 217-243.

Punt, A.E., Butterworth, D.S., de Moor, C.L., De Oliveira, J.A.A., and Haddon, M. 2016. Management strategy evaluation: best practices. Fish Fish. 17: 303-334.

R Core Team. 2016. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available from https://www.R-project.org/ [accessed 1 January 2019].

Rice, J. 2018. A comparison of candidate management procedures for Atlantic bluefin tuna. ICCAT SCRS/2018.

Richardson, D.E., Marancik, K.E., Guyon, J.R., Lutcavage, M.E., Galuardi, B., Lam, C.H., Walsh, H.J., Wildes, S., Yates, D.A., and Hare, J.A. 2016. Discovery of a spawning ground reveals diverse migration strategies in Atlantic bluefin tuna (Thunnus thynnus). PNAS 113(12): 3299-3304.

Rodriguez-Marin, E., Ortiz, M., de Urbina, J.M.O., Quelle, P., Walter, J., Abid, N., Addis, P., Alot, E., Andrushchenko, I., Deguara, S., Di Natale, A., Gatt, M., Golet, W., Karakulak, S., Kimoto, A., Macias, D., Saber, S., Neves Santos, M., and Zarrad, R. 2015. Atlantic bluefin tuna (*Thunnus thynnus*) biometrics and condition. PLoS ONE 10(10):e0141478. doi:10.1371/journal.pone.0141478.

Sainsbury, K.J., Punt, A.E., and Smith, A.D.M. 2000. Design of operational management strategies for achieving fishery ecosystem objectives. ICES J. Mar. Sci. 57: 731-741.

Ying, Y., Chen, Y., Lin, L., and Gao, T. 2011. Risks of ignoring fish population spatial structure in fisheries management. Can. J. Fish. Aquat. Sci. 68: 2101-2120.

Zarrad, R., Walter, J., and Lauretta, M. 2018. VPA-2BOX model diagnostics used in the 2014 assessment of eastern Atlantic bluefin tuna. ICCAT Collect. Vol. Sci. Pap. 74(6): 2809-2831.

Table 1. Operating model specifications. For further details on the Atlantic bluefin tuna operating model structure see Kerr *et al.* (2016, 2018) and Morse (2018).

	West	**East**
	Stocks	
Geographic zones in stock area (Fig. 1)	1-3	4-7
Number of fleets (historic)	17	10
Number of fleets (projected)	6	5
	Populations	
Age classes	1-16+	1-10+
Natural mortality-at-age vector	0.38, 0.30, 0.24, 0.20, 0.18, 0.16, 0.14, 0.13, 0.12, 0.11, 0.11, 0.11, 0.10, 0.10, 0.10	0.38, 0.30, 0.24, 0.20, 0.18, 0.16, 0.14, 0.13, 0.12, 0.10
Spawning fraction-at-age vector	0, 0, 0, 0, 0, 0.001, 0.007, 0.039, 0.186, 0.563, 0.879, 0.976, 0.996, 0.999, 1, 1 (Porch and Hanke 2018)	0, 0, 0.25, 0.5, 1, 1, 1, 1, 1, 1 (ICCAT 1999)
Growth parameters (length-at-age)	Richards model (Ailloud *et al.* 2017): L_1 = 33.0, L_2 = 270.6, p = -0.12, A_1 = 0, A_2 = 34, K = 0.22	Von Bertalanffy model (Cort 1991): K = 0.093, L_∞ = 319, t_0 = -0.97
Growth equation (length-weight; Rodriguez-Marin *et al.* 2015)	$W_a = 0.0000177054 * L_a^{3.001251847}$	$W_a = 0.0000350801 * L_a^{2.878451}$
Hockey stick stock-recruitment parameters	R_{max} = 142,883 SSB* = 4530	R_{max} = 2,683,004 SSB* = 86,437

Table 2. Indices included in the 2017 stock assessment of Atlantic bluefin tuna (ICCAT 2018a) and indications as to whether they were included in the historical or projected operating model (OM) for this MSE.

Index	Historical OM	Projected OM
CAN_Combined_RR		
CAN_GSL_Acoustic	X	X
US_RR<145	X	
US_RR_66-144	X	X
US_RR_115-144	X	X
US_RR_145-177		
US_RR>195	X	
US_RR>195_COMB		
US_RR>177		
JLL_AREA_2_(WEST)	X	
LARVAL_ZERO_INFLATED	X	X
GOM_PLL_1-6	X	X
JLL_GOM	X	
TAGGING		
JLL_RECENT	X	X
MOR_SP_TP	X	
MOR_POR_TP	X	X
JPN_LL_EastMed	X	
JPN_LL1_NEA	X	
JPN_LL2_NEA	X	X
SP_BB1	X	
SP_BB2	X	X
FR_AER1	X	
FR_AER2	X	X
WMED_LARV	X	X

Table 3. MSE performance metrics and their definitions (Carruthers 2019).

Performance metric	Definition
Average annual variation in yield	$AAVY = \frac{1}{Y-y}\sum_{y}^{Y}\lvert C_y - C_{y-1}\rvert / C_{y-1}$ where y and Y are the first and last projected years respectively
Annual average yield	Mean true yield (from the operating model) over the first and second 10-year periods of management procedure (MP) application (2016-2025, 2026-2035)
SSB depletion relative to deterministic equilibrium at F=0	SSB depletion calculated relative to the deterministic equilibrium in the absence of catches after 10 and 20 years of MP application. $SSB_{y=10}/SSB^{*}_{F=0}$ and $SSB_{y=20}/SSB^{*}_{F=0}$, where $SSB^{*}_{F=0}$ is the deterministic equilibrium SSB at F=0. This latter term can be calculated using SPR (SSB/R) at F=0 and equilibrium R*: $SSB^{*}_{F=0} = SPR_{F=0}R^{*}$ where R* is the mean R for all years since MP application.
Lowest SSB depletion	Lowest SSB depletion, $SSB/SSB^{*}_{F=0}$, over all years of MP application $\frac{\min(SSB_{MP,y})}{SSB^{*}_{F=0}}$
SSB depletion relative to zero catch trajectory	SSB depletion after the maximum number of projected years relative to the trajectory that would have occurred had no catches been taken over the full period for which MP application is being considered: $\frac{SSB_{MP,ymax}}{SSB_{F=0,ymax}}$ where $SSB_{MP,ymax}$ is the true SSB in the last projection year resulting from the MP, and $SSB_{F=0,ymax}$ is the true SSB in the last projection year with no fishing (since year 0), where years are measured from the start of projections (i.e., year 0 is the last year of the historical period and year 1 is the first projected year).
Lowest SSB depletion relative to zero catch trajectory	Lowest SSB depletion over all years of MP application relative to the zero catch trajectory $\frac{\min(SSB_{MP,y})}{SSB_{F=0,ymax}}$

Table 4. Mean true (operating model, OM) population and stock views and estimated $F_{0.1}$ for each assessment period (1 = historic, 2-8 = projected).

	West			*East*		
Period	**OM population**	**OM stock**	**Estimated**	**OM population**	**OM stock**	**Estimated**
1	0.114	0.105	0.170	0.161	0.090	0.202
2	0.118	0.113	0.112	0.106	0.098	0.129
3	0.120	0.114	0.105	0.125	0.100	0.120
4	0.119	0.114	0.119	0.122	0.100	0.104
5	0.119	0.114	0.117	0.121	0.100	0.114
6	0.120	0.114	0.111	0.120	0.100	0.121
7	0.119	0.114	0.115	0.120	0.100	0.118
8	0.119	0.114	0.111	0.121	0.100	0.122

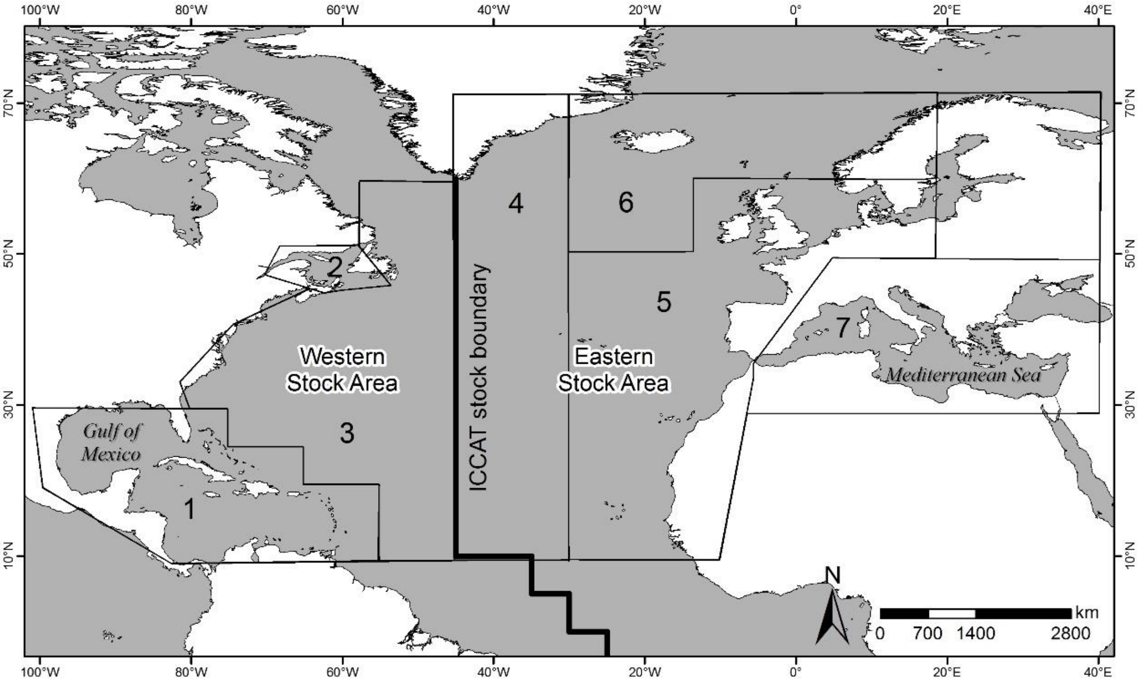


Figure 1. Spatial structure of the Atlantic bluefin tuna MSE operating model: Gulf of Mexico (zone 1), Gulf of St. Lawrence (zone 2), western Atlantic (zone 3), central Atlantic (zone 4), eastern Atlantic (zone 5), northeast Atlantic (zone 6), and Mediterranean Sea (zone 7; Kerr *et al.* 2016).

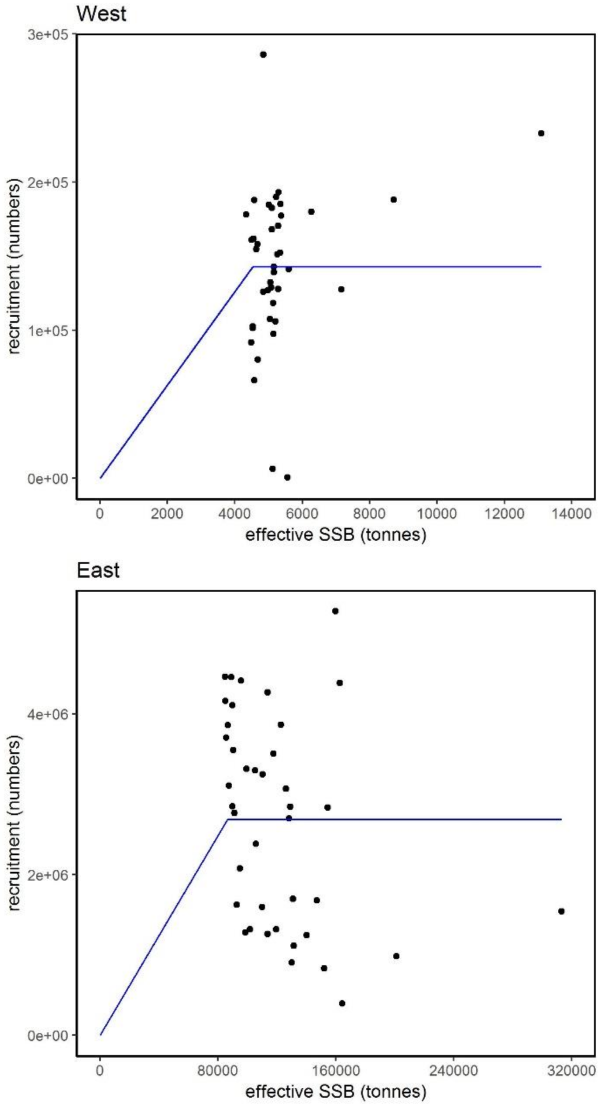
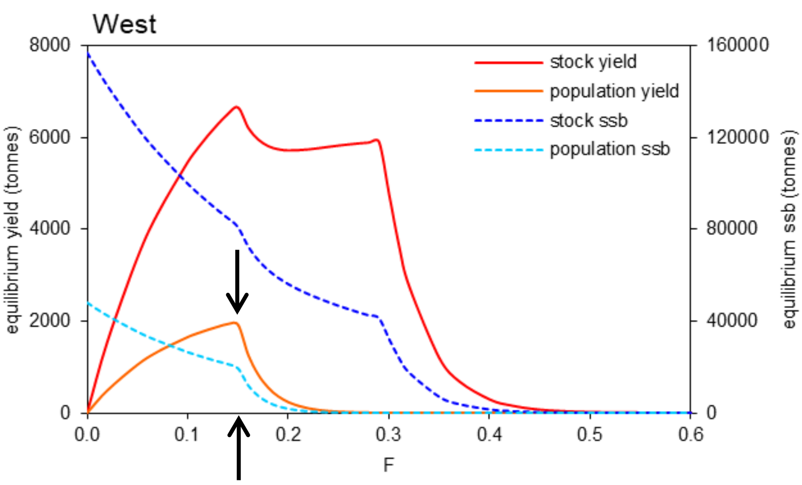


Figure 2. Hockey stick stock-recruit models used in the projected operating model. Parameters were estimated from R (ICCAT 2018a) and effective SSB (west zone 1, east zone 7) for the years 1975-2015.



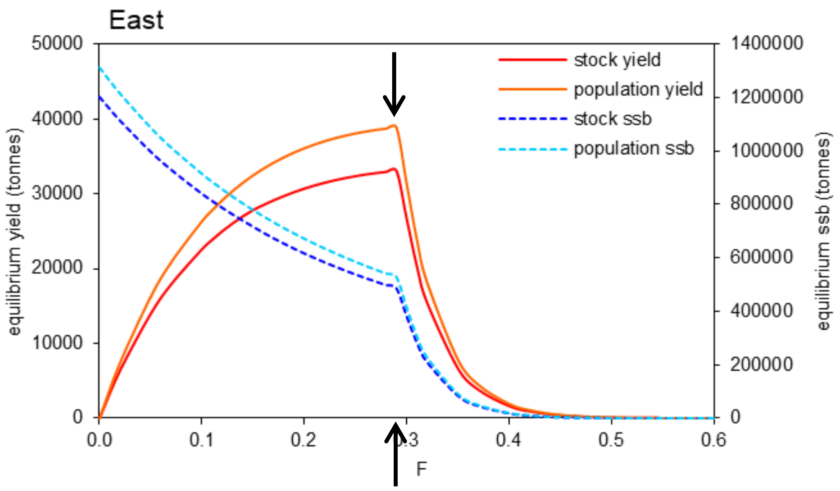
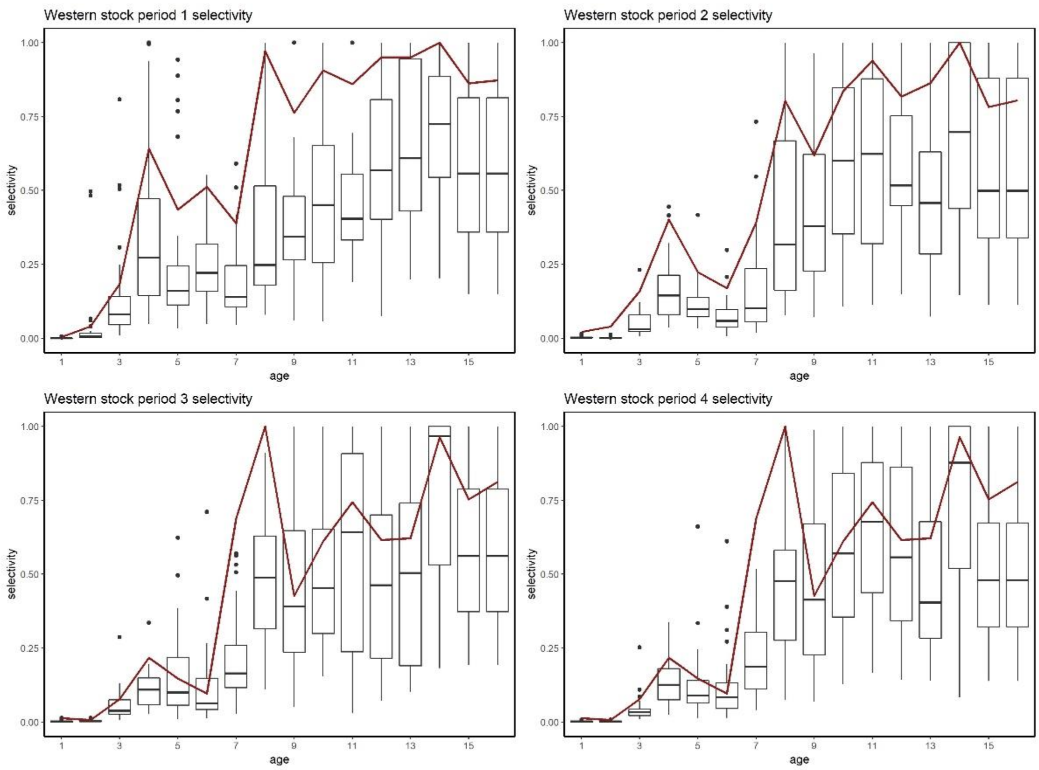


Figure 3. Long-term (150-year) deterministic equilibrium yield and SSB at a range of values for F. Black arrows indicate F_{MSY} and F_{crash}.



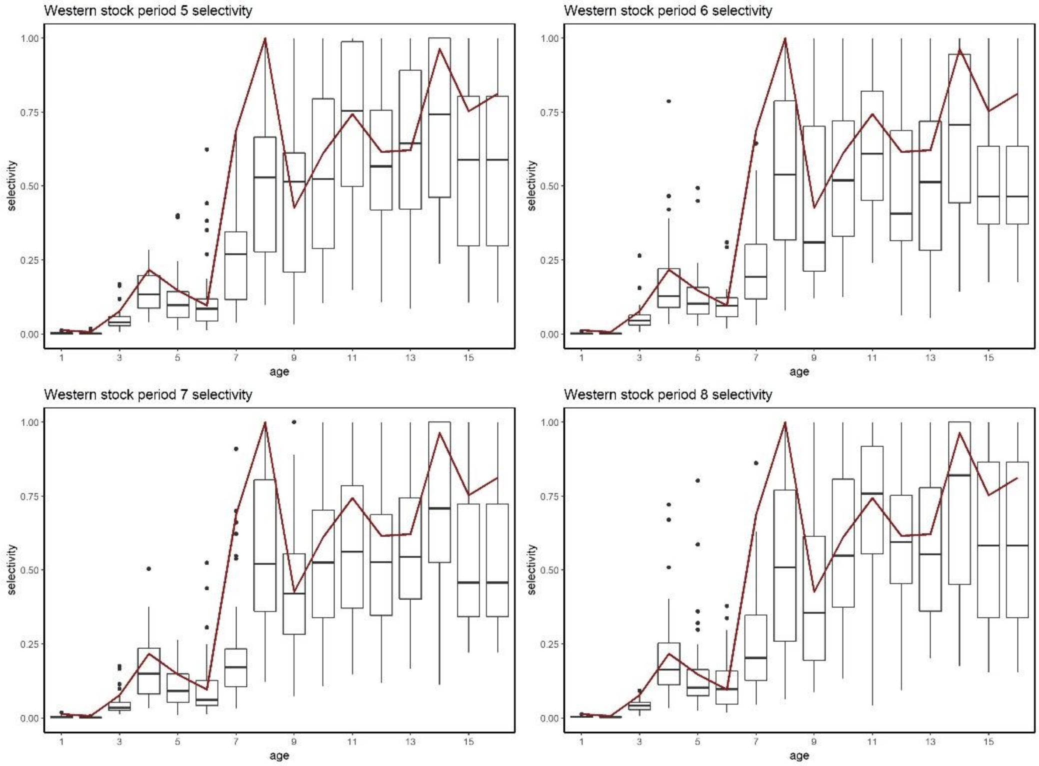
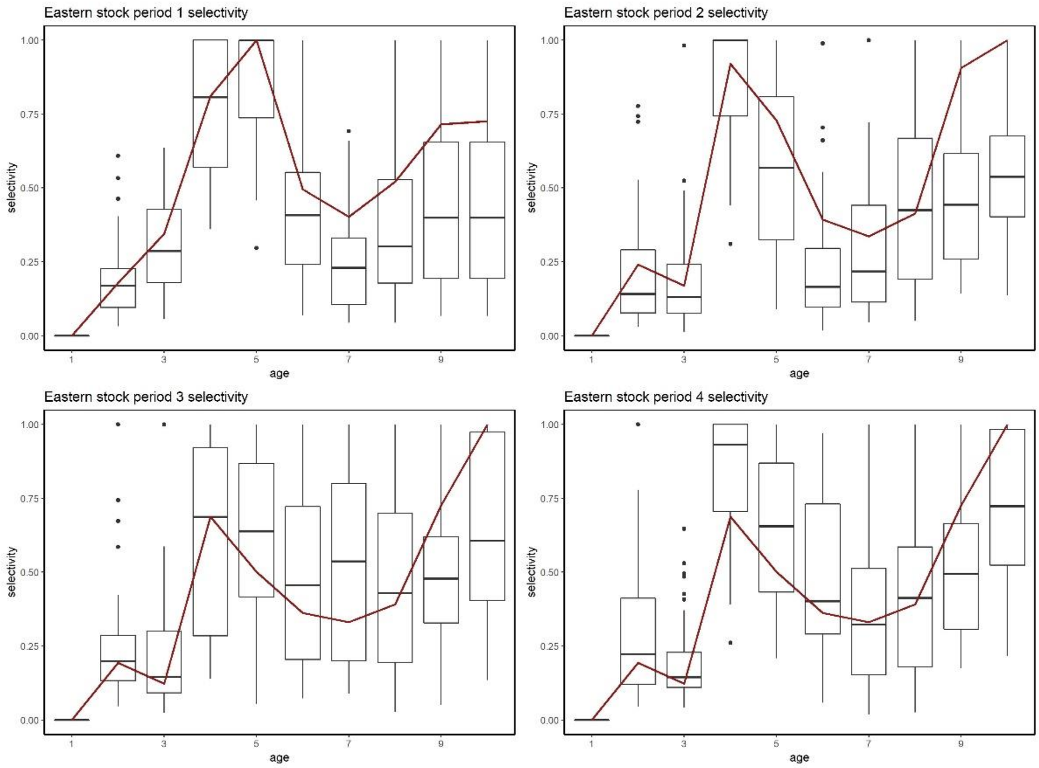


Figure 4. True (red) and stochastic estimated (boxplots of distribution of 100 MSE realizations) stock selectivities for the western stock for each of the historical period (1) and projected assessment periods (2-8).



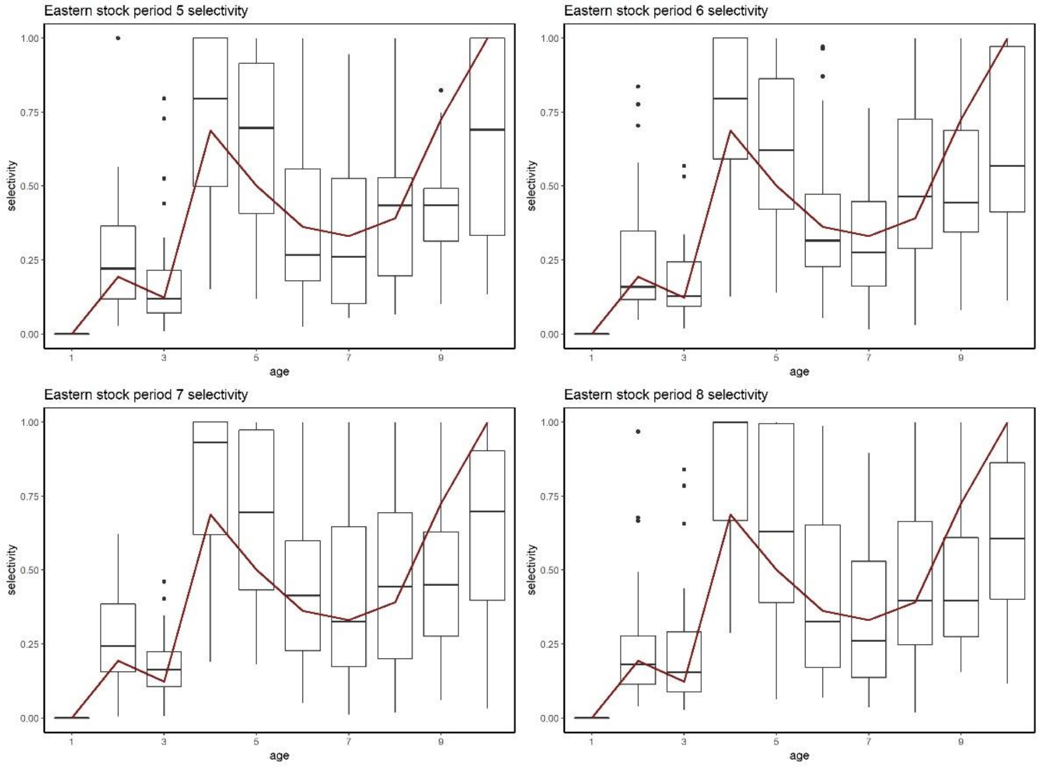
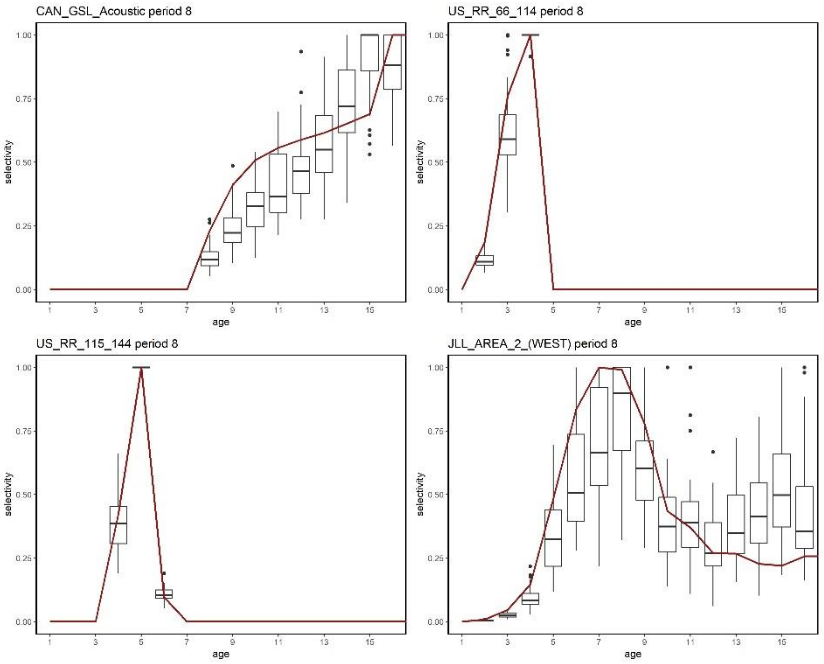


Figure 5. True (red) and stochastic estimated (boxplots of distribution of 100 MSE realizations) stock selectivities for the eastern stock for each of the historical period (1) and projected assessment periods (2-8).



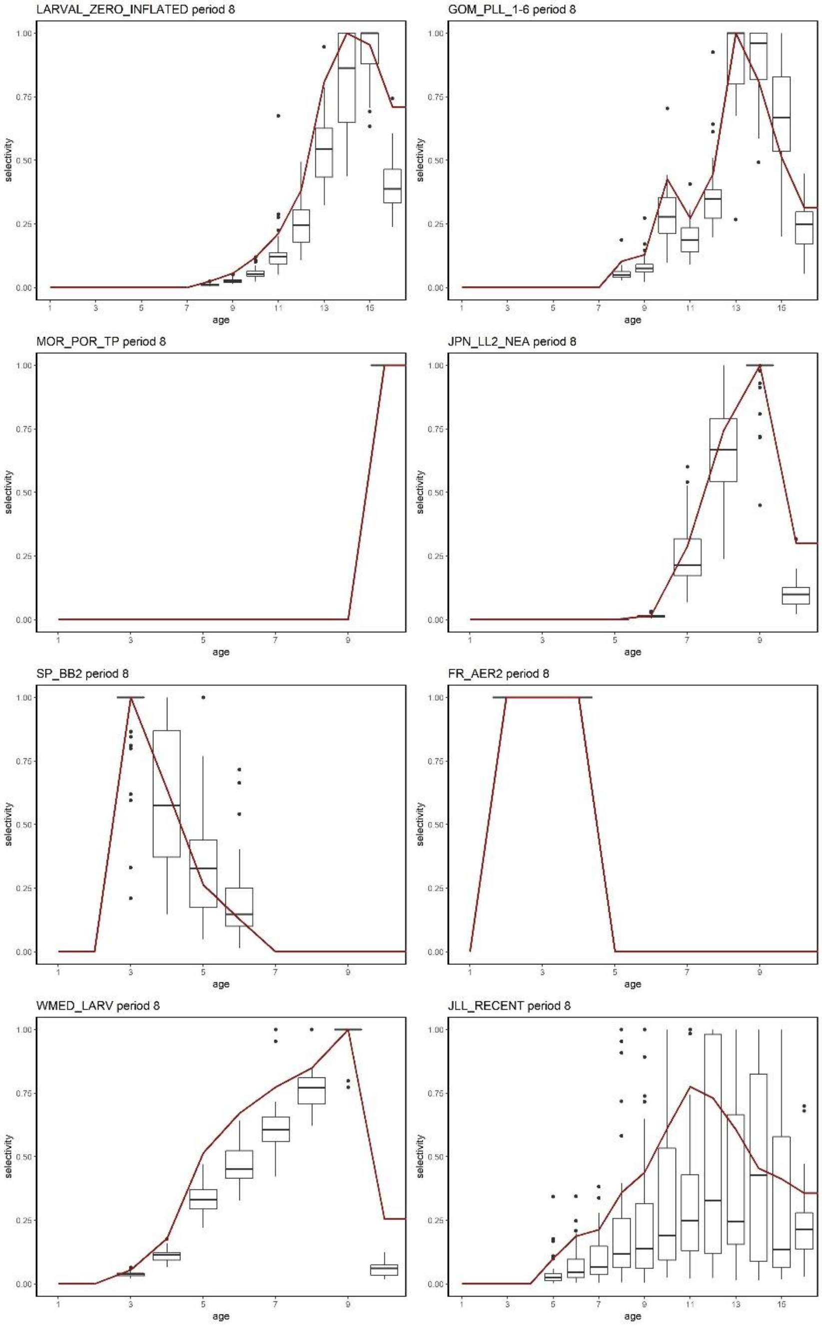


Figure 6. True (red) and stochastic estimated (boxplots of distribution of 100 MSE realizations) selectivities for indices of relative abundance at the end of projections (period 8).

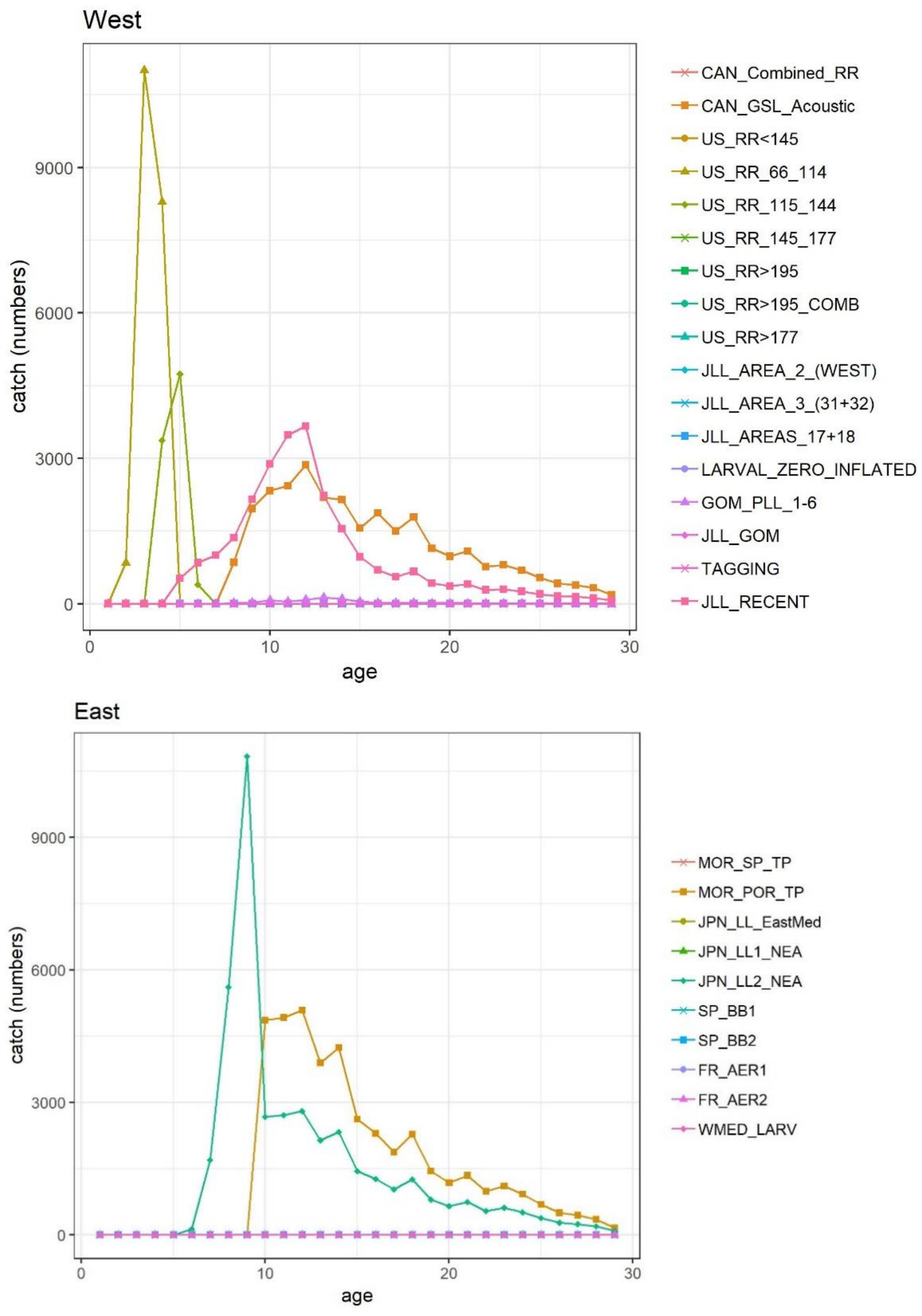


Figure 7. Deterministic age composition of catch for indices of relative abundance for the last year of the operating model historical period (2015).

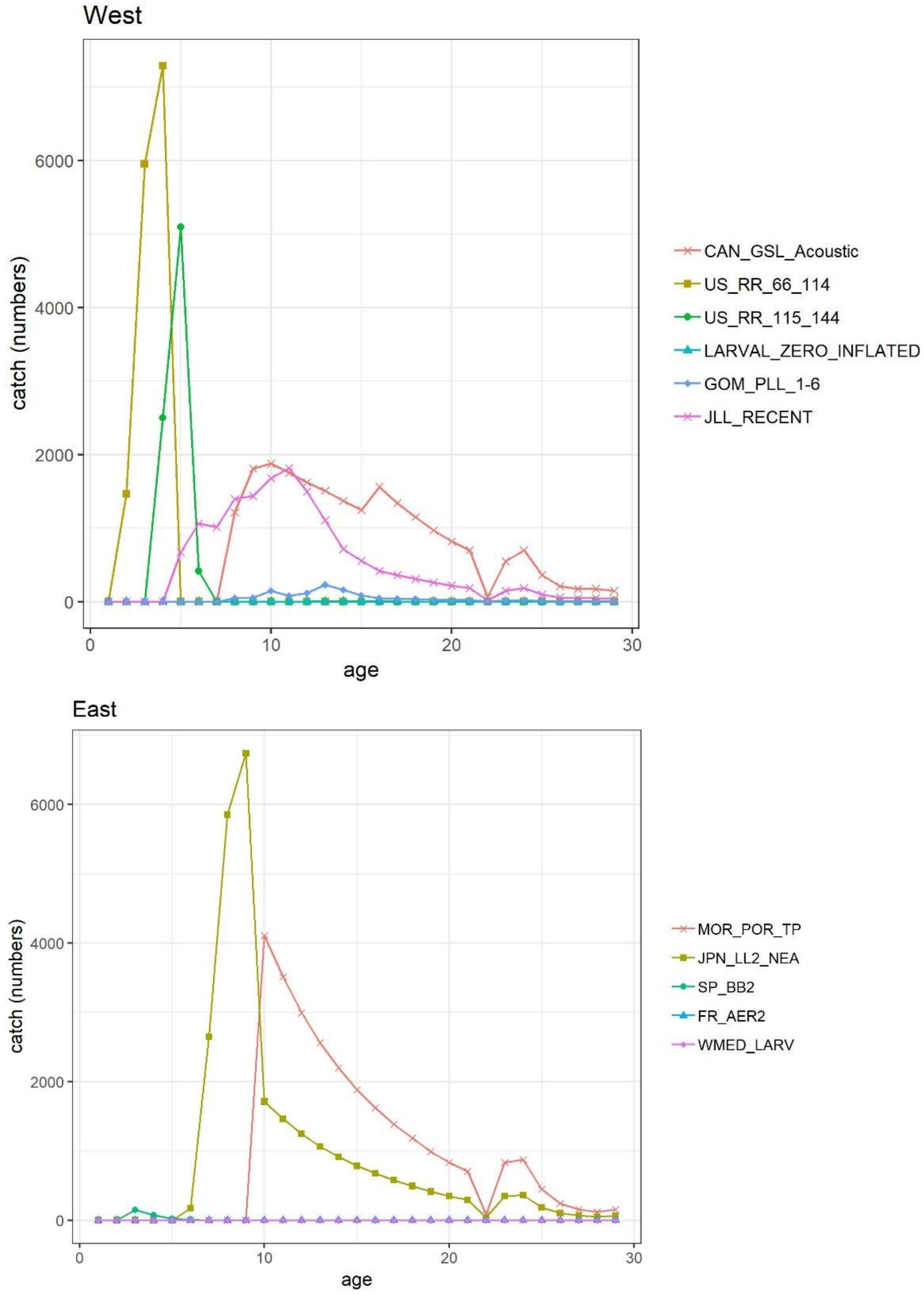


Figure 8. Deterministic age composition of catch for indices of relative abundance for the last year of the operating model projected period (2036).

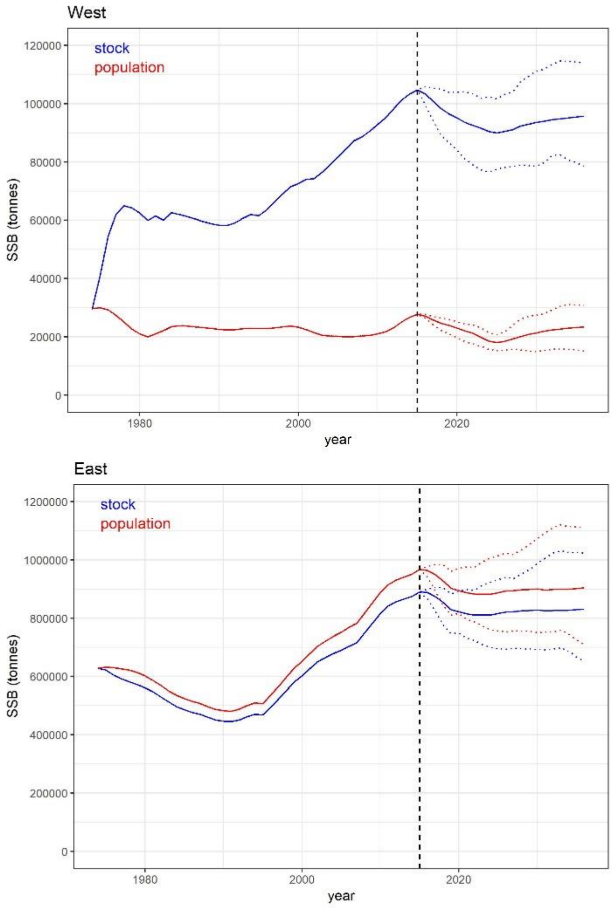


Figure 9. Time series of western and eastern population and stock SSB, with 95% confidence intervals (dotted lines) for 100 projected realizations. Dashed vertical line indicates beginning of projected period.

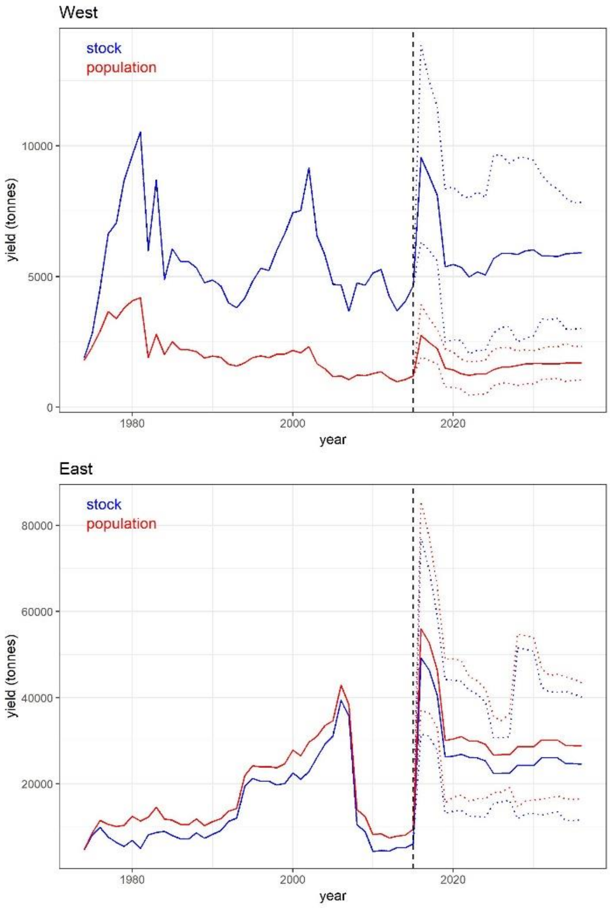


Figure 10. Time series of western and eastern population and stock yields, with 95% confidence intervals (dotted lines) for 100 projected realizations. Dashed vertical line indicates beginning of projected period.

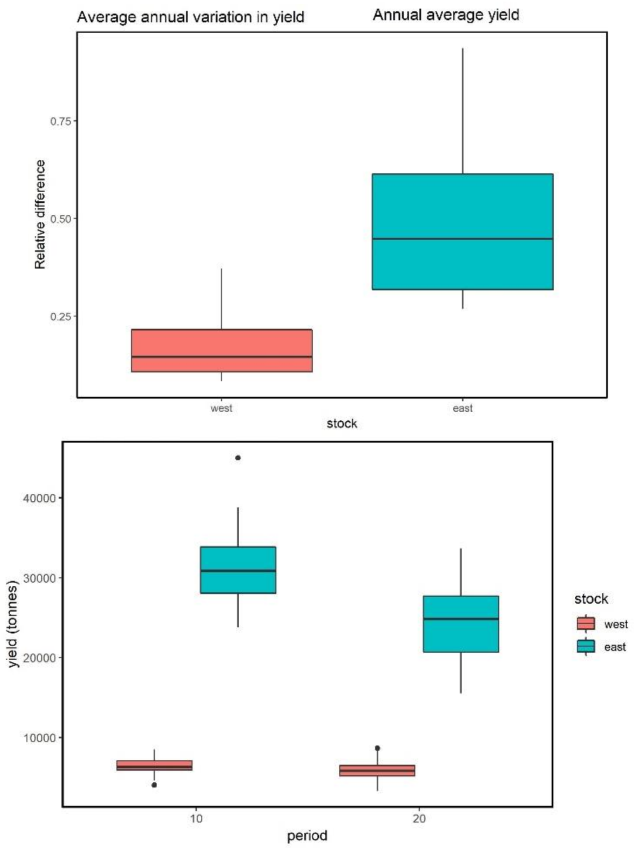


Figure 11. Average annual variation in yield over entire management procedure application (left), and annual average yield for the first and second 10-year periods of management procedure application for 100 MSE realizations.

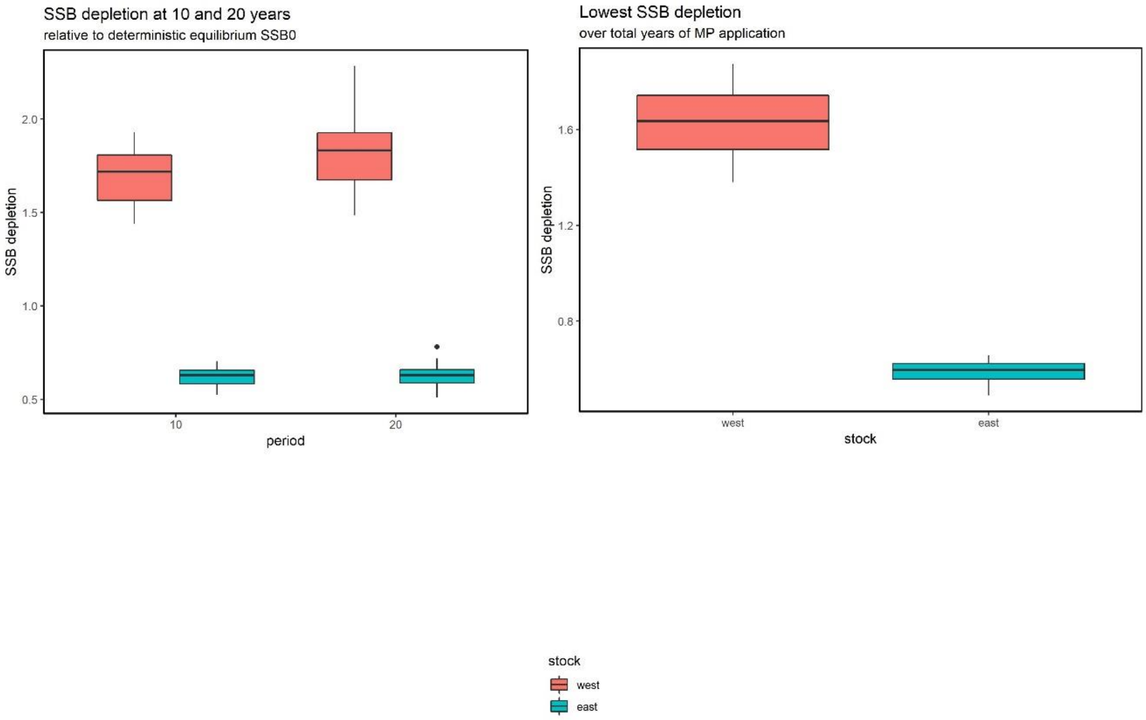


Figure 12. Stock SSB depletion relative to the deterministic equilibrium SSB_0 in the absence of catches after 10 and 20 years of management procedure (MP) application for 100 MSE realizations. Shown with the lowest SSB depletion over entire MP application.

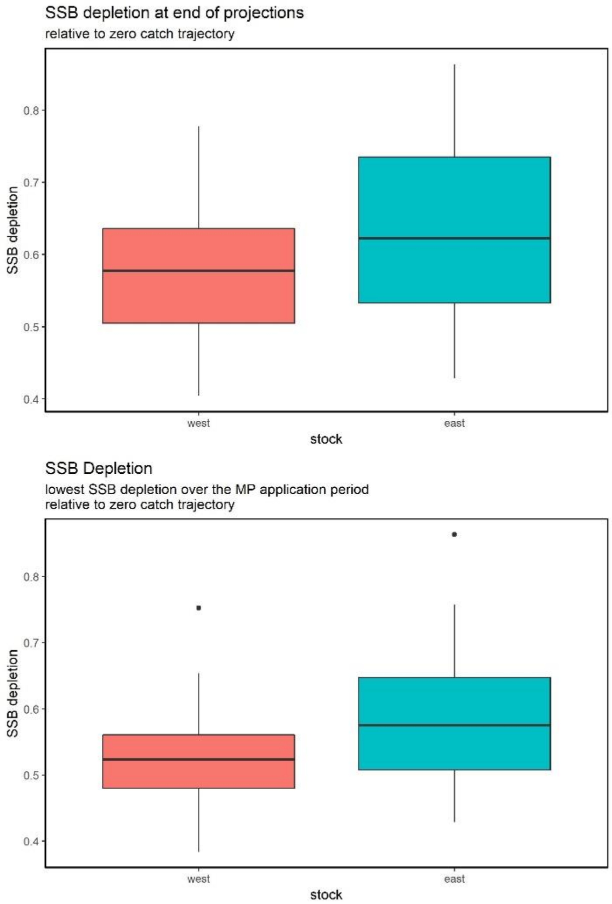


Figure 13. SSB depletion relative to the zero catch trajectory for 100 MSE realizations. Shown with the lowest SSB depletion over entire management procedure (MP) application.

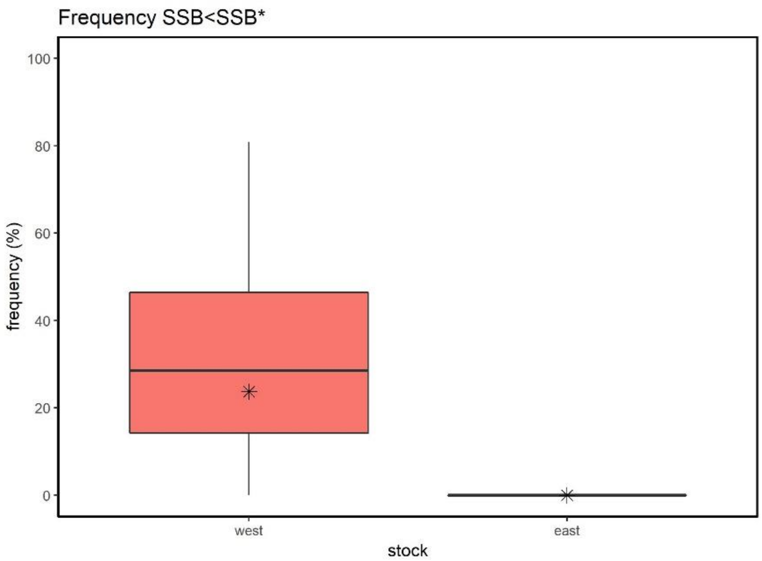


Figure 14. Frequency of SSB falling below the hockey stick hinge point, SSB*, across 100 MSE realizations with the distribution over all years (boxplots) and the final year (stars).

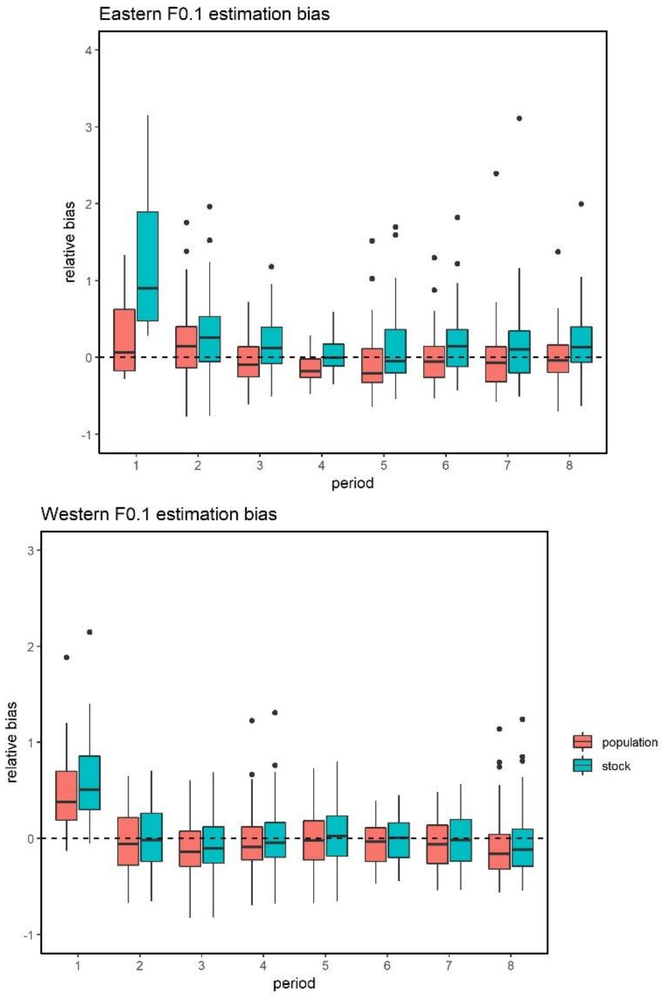


Figure 15. Estimation bias in $F_{0.1}$ for each assessment-management cycle from 100 MSE realizations (calculated for each realization as the relative error in the estimated $F_{0.1}$ relative to each of the true $F_{0.1}$ for the population and stock views, respectively).