SCRS/2017/178 Collect. Vol. Sci. Pap. ICCAT, 74(6): 3427-3441 (2018)

SIMULATING VIRTUAL POPULATION ANALYSIS OF MIXED ATLANTIC BLUEFIN TUNA STOCKS

M.R. Morse[1], L.A. Kerr[2], S.X. Cadrin[3]

SUMMARY

The purpose of this investigation was to simulation test the performance of calibrated virtual population analysis for assessing mixed Atlantic bluefin tuna stocks. Pseudodata with the typical patterns, quantity, and quality of data available for the most recent stock assessment of Atlantic bluefin tuna were generated using a previously developed operating model framework that incorporated movement and mixing between stocks conditioned on previous Atlantic bluefin tuna stock assessments. Separate eastern and western stocks were assessed using VPA-2BOX as the estimation model, and model performance was assessed by comparing results across simulations and to the stock and population views of the operating model. The estimation model was sensitive to process error (i.e., stock mixing) and measurement error, biasing estimates of spawning stock biomass, recruitment, and apical fishing mortality. The results suggest that separate virtual population analyses of eastern and western stocks accurately reflect general stock and population trends, but absolute estimates are considerably biased and may provide misleading management advice if the simulations are realistic.

RÉSUMÉ

Le but de cette étude était de tester par simulation les performances d'analyses de population virtuelle calibrées pour évaluer les stocks mixtes de thon rouge de l'Atlantique. Des pseudo-données dotées des caractéristiques typiques, quantité et qualité des données disponibles pour la plus récente évaluation du stock de thon rouge de l'Atlantique ont été générées à l'aide d'un cadre du modèle opérationnel antérieurement développé qui incorporait les déplacements et le mélange entre les stocks conditionnés aux précédentes évaluations de stocks de thon rouge de l'Atlantique. Des stocks distincts Est et Ouest ont été évalués à l'aide de VPA-2BOX comme le modèle d'estimation, et les performances des modèles ont été évaluées en comparant les résultats de différentes simulations avec la perspective du stock et de la population du modèle opérationnel. Le modèle d'estimation a été sensible à l'erreur de processus (c.-à-d. mélange des stocks) et à l'erreur de mesure, faussant les estimations de la biomasse du stock reproducteur, du recrutement et de la mortalité par pêche apicale. Les résultats suggèrent que des analyses virtuelles distinctes de la population des stocks Est et Ouest reflètent fidèlement les tendances générales des populations et des stocks, mais les estimations absolues sont considérablement biaisées et peuvent formuler un avis de gestion trompeur si les simulations sont réalistes.

RESUMEN

El propósito de esta investigación era hacer una prueba de simulación de rendimiento de los análisis de población virtual calibrados para evaluar los stocks de atún rojo del Atlántico mezclados. Se generaron pseudodatos con los patrones típicos, cantidad y calidad de los datos disponibles para la evaluación de stock de atún rojo del Atlántico más reciente, utilizando un modelo operativo previamente desarrollado que incorporaba el movimiento y la mezcla entre los stocks condicionados en previas evaluaciones del stock de atún rojo del Atlántico. Se evaluaron stocks oriental y occidental por separado utilizando VPA-2BOX como modelo de estimación, y se evaluó el rendimiento del modelo mediante la comparación de los resultados de las diferentes simulaciones con las perspectivas del stock y de la población del modelo operativo. El modelo de estimación fue sensible al error de proceso (es decir, mezcla del stock)

[1] University of Massachusetts Dartmouth, School for Marine Science and Technology, 200 Mill Rd, Fairhaven, MA 02179, USA, mmorse1@umassd.edu
[2] Gulf of Maine Research Institute, 350 Commercial St, Portland, ME 04101, USA
[3] University of Massachusetts Dartmouth, School for Marine Science and Technology, 200 Mill Rd, Fairhaven, MA 02179, USA

y al error de medición, sesgando las estimaciones de biomasa del stock, reclutamiento y mortalidad por pesca apical. Los resultados sugieren que los análisis de población virtual separados de los stock oriental y occidental reflejan con precisión las tendencias generales del stock y la población, pero las estimaciones absolutas están considerablemente sesgadas y pueden dar lugar a un asesoramiento de ordenación erróneo si las simulaciones son realistas.

KEYWORDS

Atlantic bluefin tuna, population dynamics, simulation, stock assessment, tuna fisheries

Background

Atlantic bluefin tuna (*Thunnus thynnus*) is a highly migratory species composed of two populations that spawn in the Mediterranean Sea and Gulf of Mexico, respectively, and mix extensively in the North Atlantic Ocean (NRC 1994, ICCAT 2001, Boustany *et al.* 2008, Rooker *et al.* 2014). Bluefin tuna exhibit natal philopatry, returning to their own birthing areas to spawn, which results in genetic isolation of the two separate eastern and western populations (Block *et al.* 2005). Where bluefin tuna populations overlap in the North Atlantic, primarily for feeding, stock mixing can be considerable and varies across space, time, and demographic groups (NRC 1994, Mather *et al.* 1995, Siskey *et al.* 2016).

The International Commission for the Conservation of Atlantic Tunas (ICCAT) assesses and manages Atlantic bluefin tuna as two separate eastern and western stocks separated by a boundary at the 45°W meridian (ICCAT 2014). In conducting stock assessments and assigning annual quotas, it is assumed that the level of geographic mixing between the two stocks is limited (ICCAT 1992, 1993). However, with evidence increasingly demonstrating that geographic mixing of distinct populations due to regular migrations between stock areas is significant, it is important to develop an accurate understanding about the species' life history, particularly migration patterns and spawning behavior, to correctly assess the species.

ICCAT has committed to developing new stock assessment modeling approaches for Atlantic bluefin tuna to account for more complex stock structure (ICCAT 2013b). Several stock assessments of bluefin tuna have attempted to incorporate stock structure and stock mixing, and models have been developed to reflect an expanding understanding of movement and stock mixing. Stock mixing was initially modeled using a diffusion model, which assumes that bluefin tuna from one area move to another and take on the behavior of resident fish (Porch *et al.* 1998, Cadrin and Secor 2009, Taylor *et al.* 2011). However, more recent genetic, otolith, and tagging studies (Carlsson *et al.* 2004, 2007; Block *et al.* 2005, Boustany *et al.* 2008, Walli *et al.* 2009, Rooker *et al.* 2014, Secor *et al.* 2014) provide strong evidence for stock mixing more accurately described by an overlap model, which assumes that bluefin tuna return to natal areas to spawn and fish from the two different populations do not interbreed. Stock assessment models that account for stock structure and mixing may perform better than conventional models for some applications (e.g., Porch *et al.* 1998), but not for others (e.g., Goethel *et al.* 2015). However, because past simulation analyses have suggested that mixing models do not consistently perform better than models without mixing (Porch *et al.* 1998), ICCAT has continued with separate eastern and western assessments. More recent information on stock composition and advances in model development and simulation justify renewed performance evaluation of models that account for stock mixing.

Comparative evaluations are needed to assess the relative performance of stock assessment models for Atlantic bluefin tuna. Simulation testing is considered best practice for evaluating model performance, because parameter values are known (e.g., NRC 1998, Kell *et al.* 2009, Deroba *et al.* 2015). Goethel *et al.* (2016) developed a three-tier estimation-simulation-evaluation framework for estimating spatially explicit parameters to inform spatial operating models. The estimation tier involves a spatially explicit stock assessment that ideally matches the most likely population structure, patterns of movement, and mixing using available data to provide realistic estimates of spatially explicit parameters. The simulation tier involves the development of an operating model that is conditioned on the complex spatial estimation model that represents a virtual reality for generating pseudodata. Parameter estimates from the spatially complex model are then used as "known" values in a virtual reality that represents our understanding of the system of inference. The evaluation tier involves the application of spatially-simpler management procedures to the pseudodata. The operating model can be used to simulate and evaluate various management policies or test hypotheses regarding spatial structure. In the context of Atlantic bluefin tuna, this framework can be used to evaluate the performance of estimation models and management strategies.

ICCAT conducts stock assessments of Atlantic bluefin tuna using calibrated virtual population analysis (VPA). Alternative estimation models (e.g., statistical catch-at-age, statistical catch-at-length) have produced similar results to VPA (Legault and Restrepo 1998, ICCAT 2013b, Butterworth and Rademeyer 2015). Exploratory analyses of Atlantic bluefin tuna led to the development of the VPA-2BOX program (Porch *et al.* 1995, 2001; Porch 2003), which is currently used for separate assessments of western and eastern stocks (ICCAT 2014); however, VPA-2BOX does offer the option to assess two intermixing stocks simultaneously.

The purpose of this research was to use the simulation framework to test the performance of calibrated VPA stock assessment models for mixed Atlantic bluefin tuna stocks. The research objective was to use a previously developed operating model framework that accounts for spatial structure and stock mixing (Kerr *et al.* 2014, 2016), which was modified to address decisions from the 2017 ICCAT Bluefin Tuna Data Preparatory Meeting (ICCAT 2017, Kerr *et al.* 2017), to generate pseudodata with the typical patterns, quantity, and quality of data available for the most recent stock assessment of Atlantic bluefin tuna (ICCAT 2014). Pseudodata for eastern and western stocks were analyzed separately using VPA-2BOX, and model performance was assessed by comparing results across simulations and to the operating model. The goal of this research was ultimately to demonstrate how stock mixing, as represented in the operating model, might influence Atlantic bluefin tuna stock assessment results.

1. Methodology

The operating model developed by Kerr *et al.* (2017) includes two spawning populations, with the eastern population originating in the Mediterranean Sea and the western population originating in the Gulf of Mexico. The operating model is age structured (ages 1-29) and simulates movement of fish across seven geographic zones (**Figure 1**) and over four seasonal quarters (quarter 1 = spring). The overlap model assumes that bluefin tuna from one area move to another, but return to their natal area to spawn. The operating model was simulated over the 44-year time span (1970-2013) common to both eastern and western stock assessments. The 1970-2013 period was chosen to reflect the western stock assessment period, and stock identity of earlier fishery production is uncertain. The operating model was conditioned based on VPAs for eastern and western bluefin tuna stocks (ICCAT 2014, Zarrad *et al.* 2017), the recent Report of the 2017 ICCAT Bluefin Tuna Data Preparatory Meeting (ICCAT 2017), and fishery-independent estimates of movement (Galuardi *et al.* 2017). The VPAs had the same settings as the 2014 western assessment (ICCAT 2014) and the revised settings recommended by Zarrad *et al.* (2017) for the eastern assessment, and both were revised to reflect the natural mortality and maturity assumptions recommended by the data preparatory meeting (ICCAT 2017). Operating model results were provided by geographic zone, season, and population-of-origin, and results were aggregated to annual time steps and stock area to derive VPA-2BOX input data. (See Kerr *et al.* (2017) for a detailed description of the operating model.)

In order to generate pseudodata of the typical pattern, type, and quantity available for Atlantic bluefin tuna, an observation model was applied to the data generated by the operating model to derive catch-at-age, partial catch-at-age, and indices of abundance for both the eastern and western stocks. The same indices of abundance as those used in the 2014 ICCAT assessments were used. For stochastic simulations, measurement error was assumed and generated randomly using the *rnorm( )* function in R for each observation according to a normal distribution for lognormally-distributed catch-at-age, partial catch-at-age, and indices of abundance.

Pseudodata were generated with measurement error. All calculations were coded and scripts run in R version 3.3.2. For more detail on the derivation of the following equations, see Kerr *et al.* (2017).

$$I_{y,g} = \left(Q_g \sum_{a=1}^{a=29} S_{a,g} N_{y,a}\right) * e^{\varepsilon_{y,g}}$$

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

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

Eqn. 1[4]

Eqn. 2

Eqn. 3

Parameter descriptions

$I_{y,g}$	Index value (e.g., catch per unit effort) of a given fleet by year
Q_g	Catchability of a given fleet (time invariant)
$S_{a,g}$	Selectivity of fish at age by fleet
$N_{a,q,y,z,p}$	Number of fish by age, quarter, year, zone, and population
$C_{y,a,g}$	Fleet-specific partial catch-at-age by year
$E_{y,g}$	Fleet-specific effort by year
$M_{a,q,p}$	Quarterly natural mortality by age and population
$C_{y,a}$	Catch-at-age by year
$F_{y,a,q,z}$	Quarterly zone-specific fishing mortality by year

For the indices of abundance, abundance was assumed to be at the beginning of the third quarter (fall), except for the indices that measured relative abundance in spawning areas (Mediterranean Sea or Gulf of Mexico), in which case the abundance was taken from the beginning of the first quarter (spring) to reflect the spawning period.

The partial catches-at-age for all but two indices were derived from the operating model using Eqn. 2. The maturity vector was used as the partial catch-at-age for the larval index (as in Sensitivity Run 22 in the 2014 assessment) and the partial catch-at-age from the 2014 assessment was used for the tagging index, in which all fish ages 1 to 3 are equally vulnerable to the fishery (ICCAT 2014).

The error terms for the indices of abundance were normally distributed:

$$\varepsilon \sim N(0, \sigma)$$

The standard deviation, σ, of the error term for each index of abundance was calculated as the root mean square error:

$$\sigma = \sqrt{\frac{1}{Y}\sum_{y}^{Y}[\ln r_i - \ln \hat{r}_i]^2}$$

The observed, r_i, and predicted, $\hat{r}_i$, residuals were obtained from Table 5 of the results file of the revised VPAs as the untransformed observed and predicted residual values.

The coefficients of variation for the indices of abundance were derived directly from the input data files from the revised VPAs. The coefficients of variation for a given index were averaged across all years, and the resulting value was used as the coefficient of variation for all years of the index.

The error terms for the catch-at-age and partial catch-at-age were normally distributed:

$$\varepsilon \sim N(0, \sigma)$$

The standard deviation, σ, of the error term for each observation was derived as the mean squared log-residual of catch-at-age data from an exploratory age structured assessment program (ASAP; Maguire *et al.*, in press) analysis of 2014 Atlantic bluefin tuna stock assessment input data:

$$\sigma = \frac{1}{A * Y}\sum_{a}^{A}\sum_{y}^{Y}\left[\ln\left({C_{a,y}}\Big/{\hat{p}_{a,y} * \sum_a^A C_{a,y}}\right)\right]^2$$

[4] Indices measured as biomass also multiplied the selectivity and abundance by the weight-at-age, W_a.

where A is the maximum age class, Y is the last year, and $\hat{p}$ is the predicted age composition. This yielded a value of $\sigma = 0.72$, which was used to generate random error according to a normal distribution for the catch-at-age and partial catch-at-age for both eastern and western stocks.

Because the operating model produced age-based data for ages 1 to 29, and the eastern and western Atlantic bluefin tuna stock assessments consider ages 10 and 16 as plus groups, respectively, the catch-at-age and partial catch-at-age data were summed across all ages from the plus group to age 29. This produced catch series for ages 1 to 10+ for the eastern stock and 1 to 16+ for the western stock.

For each simulation, the pseudodata were transformed into the proper format required by the VPA-2BOX program, which consists of three ASCII text files: control file, data file, and parameter specification file (see Porch 2003 for formatting requirements). This formatting was automated using functions generated with R script (T. Rouyer, personal communication, 28 May 2017).

The VPA-2BOX program, version 4.01 (Porch *et al.* 2001), was used to run the stock assessment simulations using the same methods as the most recent ICCAT Atlantic bluefin tuna benchmark assessment of the western stock (ICCAT 2014) and as the Zarrad *et al.* (2017) assessment of the eastern stock. Model settings and parameters were retained, with the following exceptions:

1. In 2014, ICCAT conducted stock assessments for eastern Atlantic bluefin tuna for the years 1950 to 2013 and for western Atlantic bluefin tuna for the years 1970 to 2013. However, the operating model required conditioning information for both stocks, so the eastern bluefin tuna simulations were also limited to the period 1970-2013.

2. The maturity and natural mortality vectors were based on decisions made by the 2017 ICCAT Bluefin Tuna Data Preparatory Meeting (ICCAT 2017). The same vectors were used for both the eastern and western VPAs.

3. The weight-at-age information used in the "index weight information" and "fecundity information" sections of the VPA-2BOX input data files for both the eastern and western stock assessments was informed by the Richards growth model adopted by the 2017 ICCAT Bluefin Tuna Data Preparatory Meeting (Ailloud *et al.* 2017, ICCAT 2017). A single length-weight relationship for the Atlantic (ICCAT 2013a) was used to derive weight-at-age from length-at-age. This same weight-at-age information was used in the operating model, but it had to be converted from metric tons to kilograms for the estimation model.

4. The method suggested by Zarrad *et al.* (2017) to derive the F-ratio was adopted for eastern and western VPAs. The method estimates the terminal fishing mortality as a random deviation from the previous constant parameter or as set equal to the value of the closest previous estimated parameter in five-year blocks.

The VPA-2BOX program executable was run using a shell function in R, and VPA results files were read back into R for evaluation. The first run of the estimation model for each stock used "perfect" data (i.e., with no measurement error), and 50 simulations with measurement error were run for each stock as an initial demonstration for review by ICCAT. Model performance was assessed based on accuracy, precision, and bias of the estimation model results for spawning stock biomass, recruitment, and apical fishing mortality with the operating model outputs. Operating model outputs for spawning stock biomass were presented both as "stock view," referring to the separate eastern and western stocks separated by the 45°W meridian as defined by ICCAT, and as "population view," referring to the separate genetically-distinct eastern and western subpopulations originating in their respective natal grounds.

2. Results

The simulations with no measurement error demonstrated the ability of the VPA results to accurately capture the general trends of the "true" eastern and western populations from the operating model (**Figures 2, 3, 4**). Although the eastern VPA of pseudodata without error fit the data well, reflecting the trend and magnitude of the population view of the spawning stock biomass, results did not accurately capture the trend of the first ten years (1970 to 1980), suggesting sensitivity to process error, i.e., stock mixing (**Figure 2**). The western VPA of pseudodata without error captured the trends of western spawning stock biomass well, but estimates of spawning biomass were positively biased for most of the time series (after ~1975) compared to the population view (**Figure 2**). Most recruitment estimates from both eastern and western VPAs of pseudodata without error were positively biased (**Figure 3**). The magnitude of apical fishing mortality was much lower from the eastern VPA of pseudodata without error than the operating model, and lower but to a lesser extent from the western VPA without error (**Figure 4**).

The VPAs generally overestimated the magnitude of recruitment and apical fishing mortality, and underestimated or correctly estimated spawning stock biomass. The eastern VPAs reflected the general trend of the population spawning stock biomass, but most estimates were negatively biased, indicating sensitivity to measurement error (**Figures 5 & 6, Table 1**). There was high variability in estimates of recruitment and apical fishing mortality in the eastern VPAs, but bias was relatively low (**Figures 5, 7, 8; Table 1**).

Estimates from the western VPAs reflected the general trend of the spawning stock biomass, with a slight positive bias relative to the population view (**Figures 5 & 6, Table 1**). On average, the VPAs estimated recruitment and apical fishing mortality with positive bias (**Figures 5, 7, 8; Table 1**) for the western stock, and estimates of both were highly variable.

Distributions of final objective functions suggest that most estimation models converged, but a few positive outliers suggest that some did not (**Figure 9**).

3. Discussion

Based on the simulations without measurement error, biased estimates of spawning stock biomass indicated that the western stock assessment was slightly more sensitive to stock mixing than the eastern assessment (**Figure 2**). This deterministic test was useful for demonstrating the assessment method's bias and sensitivity to process error. Because the estimation model assumed separate closed stocks, whereas the operating model incorporated fish movement and mixing, the VPA results are expected to have some estimation bias. The discrepancy between the stock and population views of the operating model spawning stock biomass for western Atlantic bluefin tuna suggests an inaccurate view of the biological resource in the west when mixing is unaccounted for. These results are similar to stock-of-origin VPAs, which demonstrated that the western fishery VPA appears to be sensitive to stock mixing, while the eastern fishery VPA is not (Cadrin *et al.* 2017). Therefore, it may be more important to consider mixing between stocks when making assessment and management decisions for the western stock.

Although VPAs of pseudodata with measurement error accurately reflected general stock trends from the operating model, the bias and relatively high estimation variability in these simulations demonstrated sensitivity to the measurement error. This result suggests that estimates of absolute stock size are highly uncertain from the current approach of separate eastern and western Atlantic VPAs, partly because of stock mixing, but mostly because of measurement error. For example, based on simulation results the eastern assessment may underestimate the population spawning stock biomass, and the western assessment may significantly overestimate the recruitment and fishing mortality. Although apical fishing mortality is the status determination criterion, differences in apical fishing mortality between the operating model and estimation model results may also involve different selectivities (e.g., apical fishing mortality at different ages).

Due to recent updates of some biological parameters and model settings used for the Atlantic bluefin tuna stock assessment, this VPA was a hybrid between the ICCAT 2014 benchmark stock assessment and the upcoming 2017 benchmark assessment. To better reflect additional updates to the Atlantic bluefin tuna stock assessment in 2017, configuration of the operating model and estimation model used in the simulations could be revised for further evaluations. We generated random measurement error for the simulations using a combination of residuals from the revised VPA-2BOX abundance index model fits and from an exploratory ASAP analysis of the 2014 assessment data (Maguire *et al.*, in press), but other methods for error generation may be substituted. More simulations (100+) would be needed for more reliable estimates of bias. The repeatability of the simulation method offers the opportunity for modification of operating model conditioning and output data structure and the use of different estimation models, such as VPA-2BOX with overlap or ASAP. Using simulation to assess the performance of VPA-2BOX with overlap might be particularly informative because the estimation model assumes stock mixing, while the current separate VPAs method does not.

4. Acknowledgements

This work was funded by the NOAA Bluefin Tuna Research Program through a series of collaborations with Clay Porch, Matt Lauretta, John Walter, and Shannon Calay. Ben Galuardi, Dave Secor, and Nathan Taylor contributed to previous versions of the operating model. Molly Lutcavage and Walt Golet collaborated on related studies. We thank John Walter for providing the VPA-2BOX files for the revised 2014 eastern stock assessment, Tristan Rouyer for providing R script for formatting VPA-2BOX input files, and J-J Maguire for sharing his ASAP analysis of the 2014 stock assessment data, which allowed us to derive the standard deviation for the catch-at-age and partial catch-at-age random error. Finally, we thank the ICCAT Atlantic-Wide Research Program for Bluefin Tuna (GBYP) for supporting Molly Morse's participation in the "Short Training Course on VPA for Atlantic bluefin tuna – Modelling Approaches."

References

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

Block, B.A., Teo, S.L.H., Walli, A., Boustany, A., Stokesbury, M.J.W., Farwell, C.J., Weng, K.C., Dewar, H., Williams, T.D. 2005. Electronic tagging and population structure of Atlantic bluefin tuna. Nature 434: 1121-1127.

Boustany, A.M., Reeb, C.A., Block, B.A. 2008. Mitochondrial DNA and electronic tagging reveal population structure of Atlantic bluefin tuna (*Thunnus thynnus*). Mar. Biol. 156: 13-24.

Butterworth, D.S. and Rademeyer, R.A. 2015. An updated statistical catch-at-length assessment for eastern Atlantic bluefin tuna. Col. Vol. Sci. Pap. ICCAT 71(4): 1639-1659.

Cadrin, S., Morse, M., Kerr, L., Secor, D., Siskey, M. 2017. Exploratory stock assessment of eastern and western stock-of-origin Atlantic bluefin tuna accounting for stock composition. ICCAT SCRS/2017/XXX.

Cadrin, S.X. and Secor, D. 2009. Accounting for spatial population structure in stock assessment: past, present, and future. pp 405-426 in Rothschild, B.J. and Beamish, R. eds. The Future of Fishery Science in North America. Springer Verlag.

Carlsson, J., McDowell, J.R., Carlsson, J.E.L., Graves, J.E. 2007. Genetic identity of YOY bluefin tuna from the eastern and western Atlantic spawning areas. Journal of Heredity 98(1): 23-28.

Carlsson, J., McDowell, J.R., Díaz-Jaimes, P., Carlsson, J.E.L., Boles, S.B., Gold, J.R., Graves, J.E. 2004. Microsatellite and mitochondrial DNA analyses of Atlantic bluefin tuna (*Thunnus thynnus thynnus*) population structure in the Mediterranean Sea. Molecular Ecology 13: 3345-3356.

Deroba, J.J., Butterworth, D.S., Methot, Jr, R.D., De Oliveira, J.A.A., Fernandez, C., Nielsen, A., Cadrin, S.X., Dickey-Collas, M., Legault, C.M., Ianelli, J., Valero, J.L., Needle, C.L., O'Malley, J.M., Chang, Y-J., Thompson, G.G., Canales, C., Swain, D.P., Miller, D.C.M., Hintzen, N.T., Bertignac, M., Ibaibarriaga, L., Silva, A., Murta, A., Kell, L.T., de Moor, C.L., Parma, A.M., Dichmont, C.M., Restrepo, V.R., Ye, Y., Jardim, E., Spencer, P.D., Hanselman, D.H., Blaylock, J., Mood, M., Hulson, P.-J.F. 2015. Simulation testing the robustness of stock assessment models to error: some results from the ICES strategic initiative on stock assessment methods. ICES J. Mar. Sci. 72(1): 19-30.

Galuardi, B., Cadrin, S.X., Arregi, I., Arrizabalaga, H., DiNatale, A., Brown, C., Lauretta, M., Lutcavage, M. 2017. Atlantic bluefin tuna area transition matrices estimated from electronic tagging and SATTAGSIM. SCRS/2017/045.

Goethel, D.R., Kerr, L.A., Cadrin, S.X. 2016. Incorporating spatial population structure into the assessment-management interface of marine resources. In "Management Science in Fisheries: An Introduction to Simulation-based Methods" (Edwards, C.T.T. and Dankel, D.J., eds.), pp 319-347. London: Routledge.

Goethel, D.R., Legault, C., Cadrin, S.X. 2015. Testing the performance of a spatially explicit tag-integrated stock assessment model of yellowtail flounder (*Limanda ferruginea*) through simulation analysis. Can. J. Fish. Aquat. Sci. 72: 1-20.

ICCAT. 1992. Report of the ICCAT Standing Committee on Research and Statistics (SCRS). Madrid, Spain. November 2 to 6, 1992.

ICCAT. 1993. Report of the ICCAT Standing Committee on Research and Statistics (SCRS). Madrid, Spain. October 25 to November 5, 1993. COM-SCRS/93/13.

ICCAT. 2001. ICCAT workshop on bluefin mixing. Madrid, Spain. September 3 to 7, 2001. Col. Vol. Sci. Pap. ICCAT 54(2): 261-352.

ICCAT. 2013a. Report of the 2013 bluefin meeting on biological parameters review. Tenerife, Spain. May 7 to 13, 2013.

ICCAT. 2013b. Report of the 2013 meeting on bluefin stock assessment methods. Gloucester, MA, USA. July 20 to 22, 2013.

ICCAT. 2014. Report of the 2014 Atlantic bluefin tuna stock assessment session. Madrid, Spain. September 22 to 27, 2014.

ICCAT. 2017. Report of the 2017 ICCAT bluefin data preparatory meeting. Madrid, Spain. March 6 to 11, 2017.

Kell, L.T., Dickey-Collas, M., Hintzen, N.T., Nash, R.D.M., Pilling, G.M., Roel, B.A. 2009. Lumpers or splitters? Evaluating recovery and management plans for metapopulations of herring. ICES J. Mar. Sci. 66: 1776–1783.

Kerr, L.A., Cadrin, S.X., Secor, D.H., Taylor, N.G. 2014. Evaluating the effect of Atlantic bluefin tuna movement on the perception of stock units. Col. Vol. Sci. Pap. ICCAT 71(4): 1660-1682.

Kerr, L.A., Cadrin, S.X., Secor, D.H., Taylor, N.G. 2016. Modeling the implications of stock mixing and life history uncertainty of Atlantic bluefin tuna. Can. J. Fish. Aquat. Sci. dx.doi.org/10.1139/cjfas-2016-0067.

Kerr, L.A., Morse, M.R., Cadrin, S.X., Galuardi, B. 2017. Application of an Atlantic bluefin tuna operating model to generate pseudodata for stock assessment testing. ICCAT SCRS/2017/177.

Legault, C.M. and Restrepo, V.R. 1998. A flexible forward age-structured assessment program. ICCAT SCRS/1998/58.

Maguire J.-J., Hanke, A., Melvin, G., Cadrin, S.X. (in press). An exploration of bluefin tuna data in the west Atlantic with ASAP. ICCAT document SCRS/2017/153. 20p.

Mather, F.J., Mason, J.M., Jones, A.C. 1995. Historical document: life history and fisheries of Atlantic bluefin tuna. US Dep. Commer., NOAA Tech. Memo. NMFS-SEFSC 370, 165 p.

NRC (National Research Council). 1994. An Assessment of Atlantic Bluefin Tuna. The National Academies Press, Washington, D.C. 148 pp.

NRC (National Research Council). 1998. Improving Fish Stock Assessments. The National Academies Press, Washington, D.C.

Porch, C.E. 2003. VPA-2BOX Version 3.01 User's Guide. NOAA Sustainable Fisheries Division Contribution SFD/2003-2004. 68 pp.

Porch, C., Kleiber, P., Turner, S., Sibert, J., Bailey, R., Cort, J.L. 1998. The efficacy of VPA models in the presence of complicated movement patterns. Col. Vol. Sci. Pap. ICCAT 50(2): 591-622.

Porch, C.E., Restrepo, V.R., Turner, S.C., Scott, G.P. 1995. Virtual population analysis of Atlantic bluefin tuna incorporating movement and tagging data. Col. Vol. Sci. Pap. ICCAT 44(1): 183-190.

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

Rooker, J.R., Arrizabalaga, H., Fraile, I., Secor, D.H., Dettman, D.L., Abid, N., Addis, P., Deguara, S., Karakulak, F.S., Kimoto, A., Sakai, O., Macías, D., Santos, M.N. 2014. Crossing the line: migratory and homing behaviors of Atlantic bluefin tuna. Mar. Ecol. Prog. Ser. 504: 265-276.

Secor, D.H., Rooker, J.R., Gahagan, B.I., Siskey, M.R., Wingate, R.W. 2014. Depressed resilience of bluefin tuna in the western Atlantic and age truncation. Conserv. Biol. 29(2): 400-408.

Siskey, M.R., Wilberg, M.J., Allman, R.J., Barnett, B.K., Secor, D.H. 2016. Forty years of fishing: changes in age structure and stock mixing in northwestern Atlantic bluefin tuna (*Thunnus thynnus*) associated with size-selective and long-term exploitation. ICES J. Mar. Sci. 73(10): 2518-2528.

Taylor N., McAllister, M., Lawson, G., Carruthers, T., Block, B. 2011. Atlantic bluefin tuna: A novel multistock spatial model for assessing population biomass. PLoS ONE 6(12): e27693.doi:10.1371/journal.pone.0027693.

Walli, A., Teo, S.L.H., Boustany, A., Farwell, C.J., Williams, T., Dewar, H., Prince, E., Block, B.A. 2009. Seasonal movements, aggregations and diving behavior of Atlantic bluefin tuna (*Thunnus thynnus*) revealed with archival tags. PLoS ONE 4(7): e6151.

Zarrad, R., Walter, J., Lauretta, M. 2016. VPA-2BOX model diagnostics used in the 2014 assessment of eastern Atlantic bluefin tuna. ICCAT SCRS/2016/037.

Table 1. Measures of relative bias (minimum annual average, maximum annual average, and average of all simulations) for results of eastern and western stock VPA simulations compared to operating model.

		Relative bias		
		Minimum annual average	*Maximum annual average*	*Average of all simulations*
Eastern stock	*Spawning stock biomass*	-33%	-12%	-26%
	Recruitment	-33%	40%	2%
	Apical fishing mortality	-56%	121%	17%
Western stock	*Spawning stock biomass*	0%	27%	13%
	Recruitment	-97%	324%	46%
	Apical fishing mortality	-60%	320%	152%

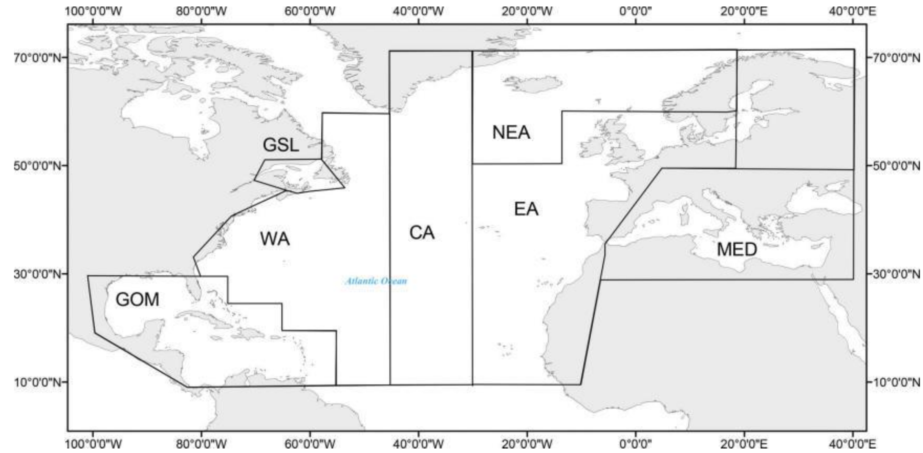

Figure 1. Geographic zones represented in the Atlantic bluefin tuna operating model: GOM, Gulf of Mexico; GSL, Gulf of St. Lawrence; WA, western Atlantic Ocean; CA, central Atlantic Ocean; EA, eastern Atlantic Ocean; NEA, northeastern Atlantic Ocean; and MED, Mediterranean Sea (Kerr *et al.* 2016).

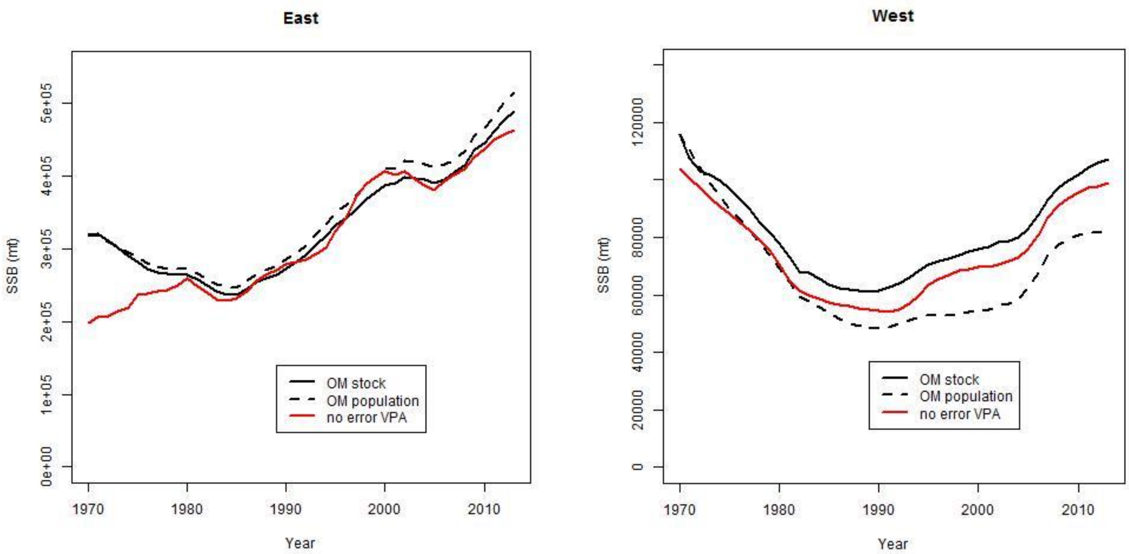

Figure 2. Spawning stock biomass time series comparing operating model (OM) stock view (solid black lines), operating model population view (dashed black lines), and results of VPAs of pseudodata without measurement error (red lines) for eastern (left) and western (right) Atlantic bluefin tuna.

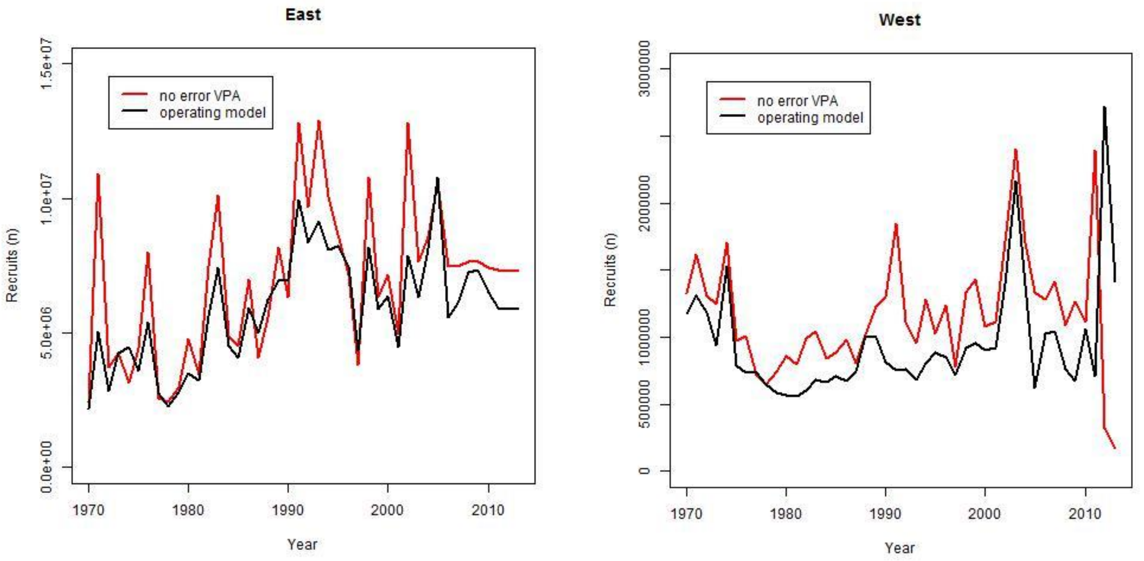

Figure 3. Recruitment time series comparing operating model (black lines) to results of VPAs of pseudodata without measurement error (red lines) for eastern (left) and western (right) Atlantic bluefin tuna.

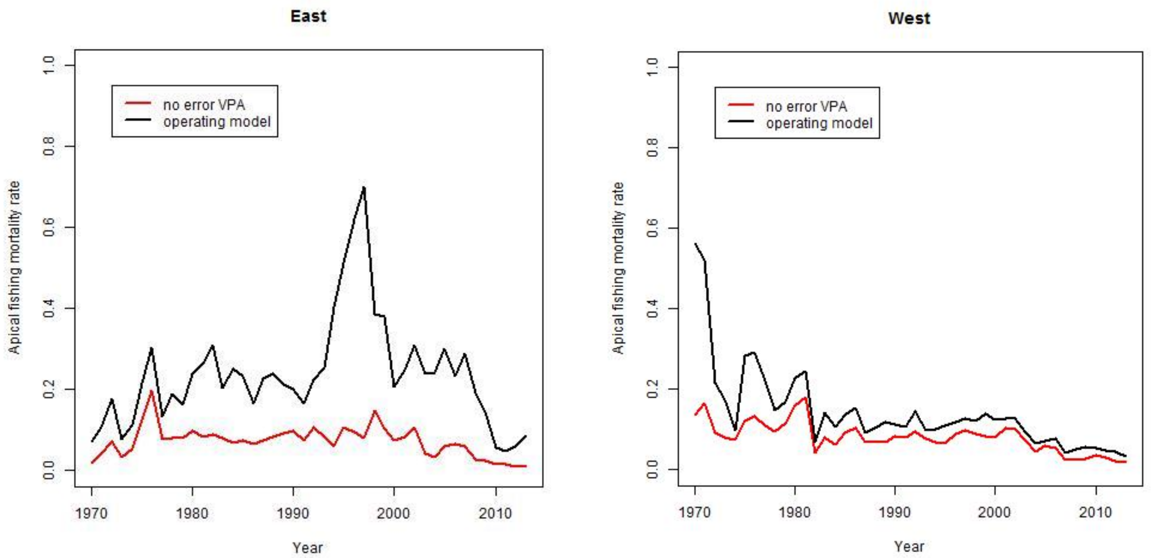

Figure 4. Apical fishing mortality time series comparing operating model (black lines) to results of VPAs of pseudodata without measurement error (red lines) for eastern (left) and western (right) Atlantic bluefin tuna.

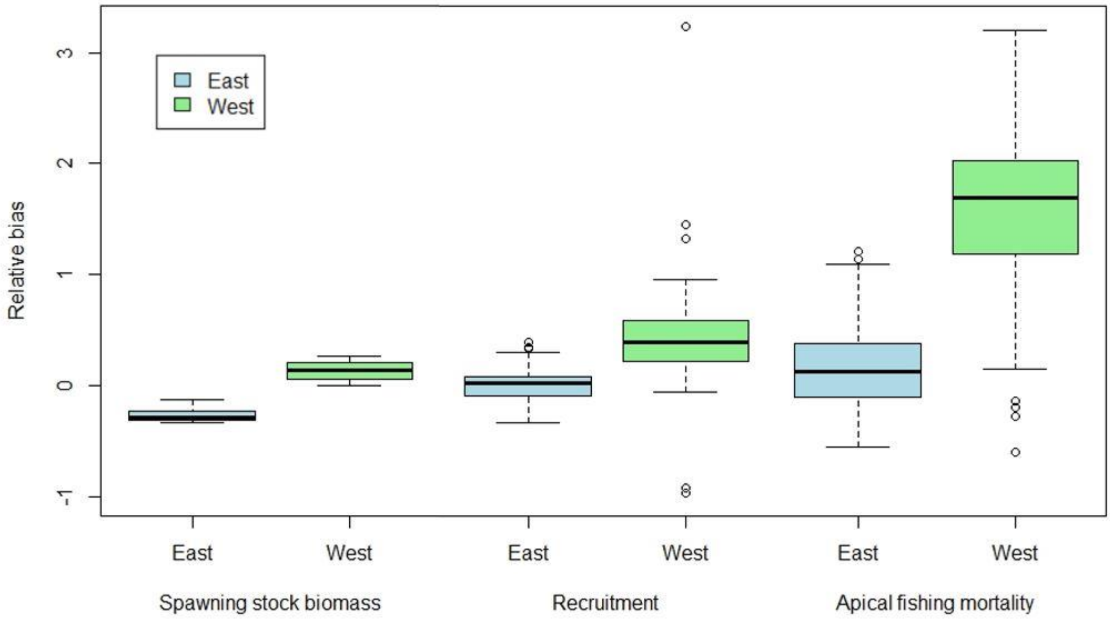

Figure 5. Annual average relative bias for spawning stock biomass, recruitment, and apical fishing mortality of the eastern and western VPA simulations compared to the operating model.

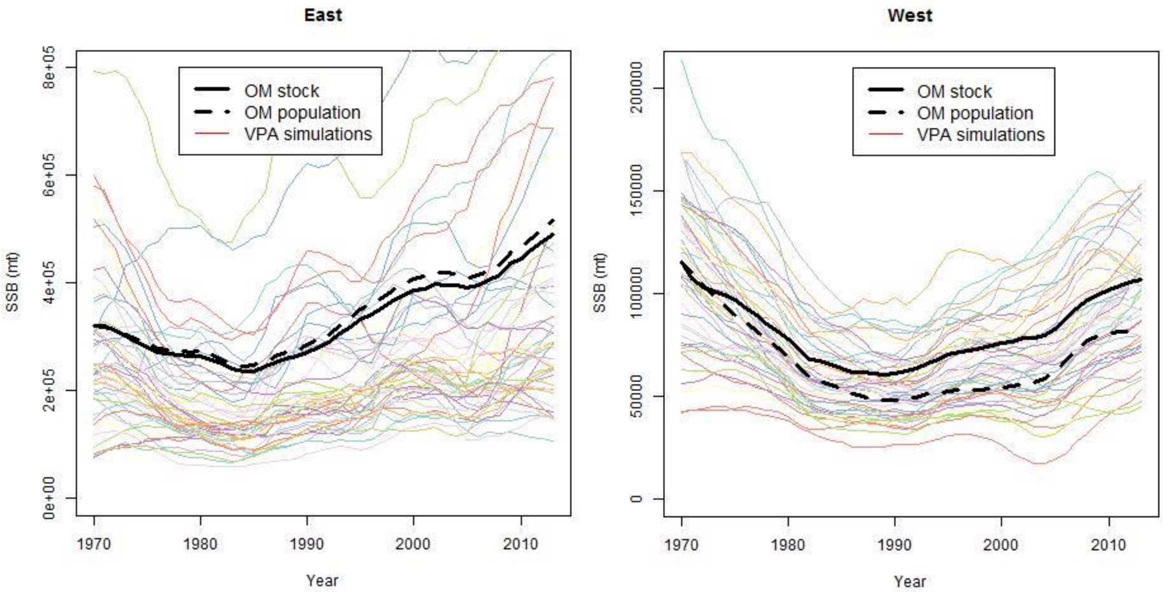

Figure 6. Spawning stock biomass time series comparing operating model (OM) stock view (solid black lines), operating model population view (dashed black lines), and VPA estimates from 50 simulations with measurement error (colored lines) for eastern (left) and western (right) Atlantic bluefin tuna.

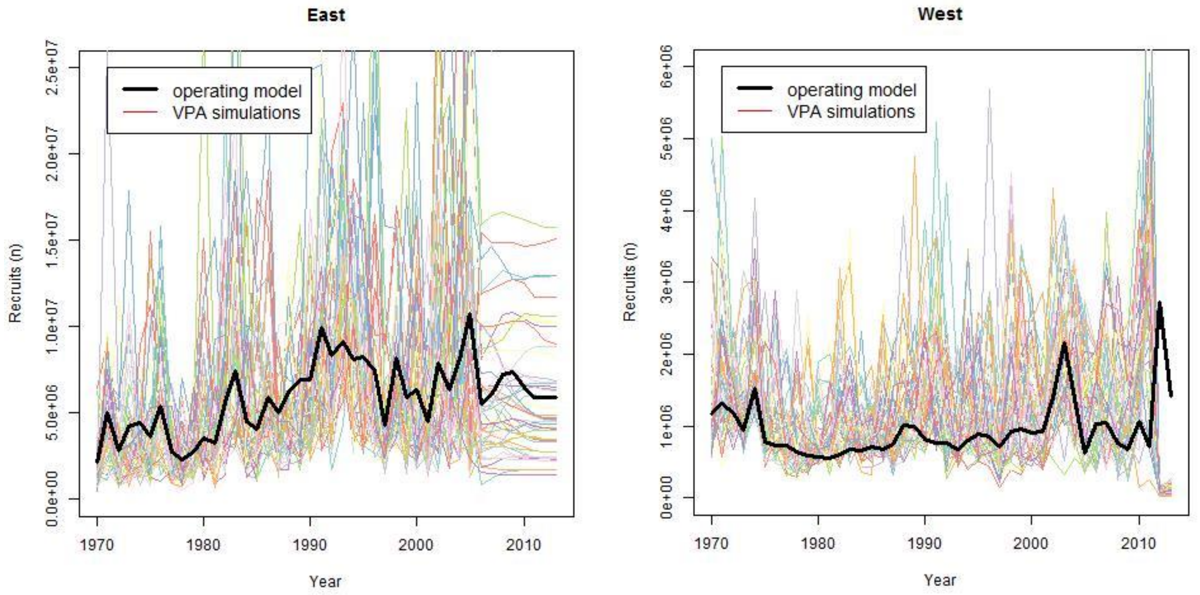

Figure 7. Recruitment time series comparing operating model (black lines) to VPA estimates from 50 simulations with measurement error (colored lines) for eastern (left) and western (right) Atlantic bluefin tuna.

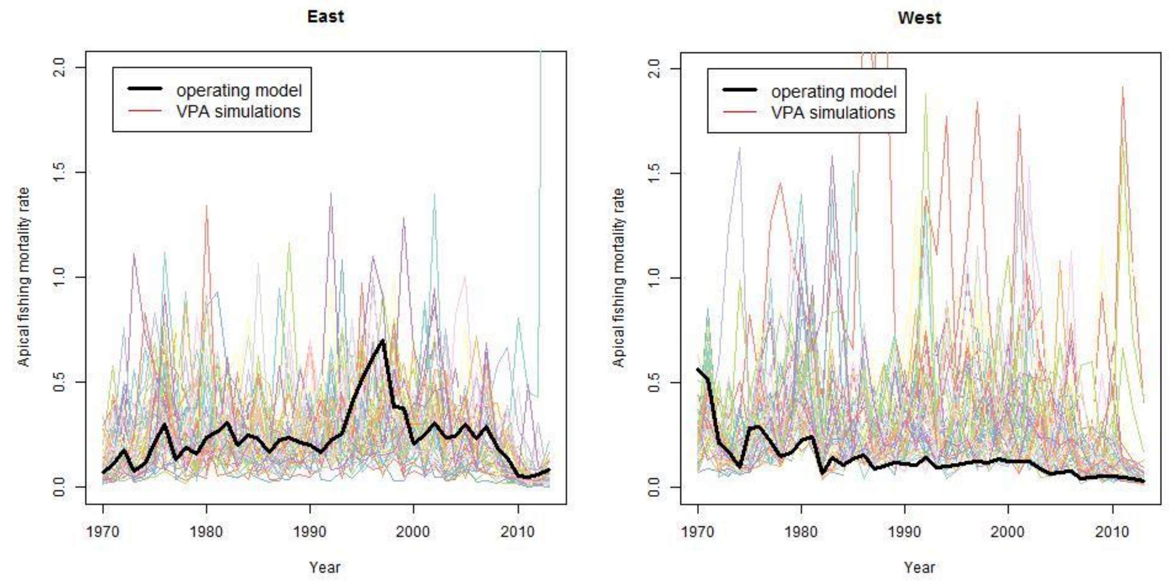

Figure 8. Apical fishing mortality time series comparing operating model (black lines) to VPA estimates from 50 simulations with measurement error (colored lines) for eastern (left) and western (right) Atlantic bluefin tuna.

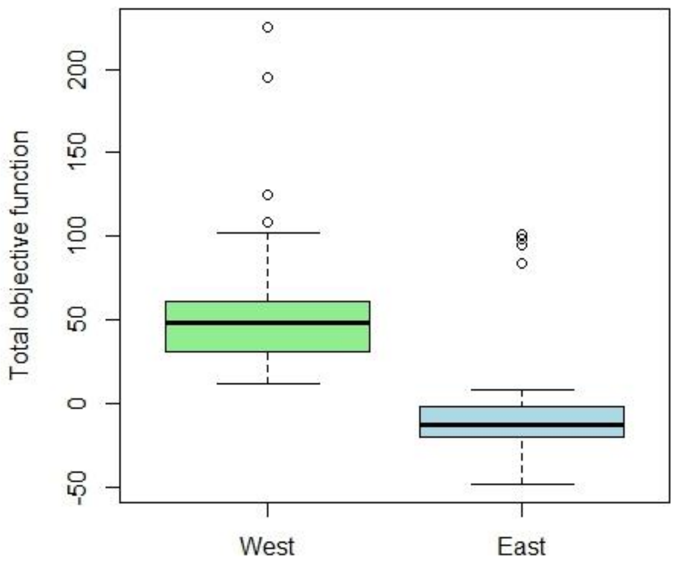

Figure 9. VPA-2BOX objective function values for 50 simulations of west and east VPAs.