 Collect. Vol. Sci. Pap. ICCAT, 77(2): 762-770 (2020)

A GLM APPROACH FOR DETERMINING THE INFLUENCE OF OPERATING MODEL FEATURES ON MANAGEMENT PROCEDURE PERFORMANCE

A.R. Hanke[1], H. Arrizabalaga [2], E. Andonegi[2] and Nick Duprey[1]

SUMMARY

GLMs were fit to the performance metrics generated by management procedures applied to models of the Atlantic Bluefin tuna fishery in a closed loop simulation. The models identified the features of the population model that accounted for the most variability in the average catch and biomass ratio over 30 years of simulated management. The variability in the performance metrics of the alternative management procedures tested was attributed to a differing set of population model features, i.e. the most influential axes of uncertainty in the population model were management procedure dependent.

RÉSUMÉ

Les GLM ont été ajustés aux mesures de performance générées par les procédures de gestion appliquées aux modèles de la pêche au thon rouge de l'Atlantique dans une simulation en boucle fermée. Les modèles ont permis d'identifier les caractéristiques du modèle de population qui représentaient la plus grande variabilité du ratio moyen entre les prises et la biomasse sur 30 ans de gestion simulée. La variabilité des mesures de performance des procédures de gestion alternatives testées a été attribuée à un ensemble différent de caractéristiques du modèle de population, c'est-à-dire que les axes d'incertitude les plus influents dans le modèle de population dépendaient de la procédure de gestion.

RESUMEN

Se ajustaron los GLM a la medición del desempeño generada por los procedimientos de ordenación aplicados a modelos de la pesquería de atún rojo del Atlántico en una simulación de círculo cerrado. Los modelos identificaron las características del modelo de población que respondía de la mayoría de la variabilidad en la ratio media de captura y biomasa durante los 30 años de la ordenación simulada. La variabilidad en la medición del desempeño de los procedimientos de ordenación alternativos probados se atribuyó a un conjunto diferente de características del modelo de población, es decir, los ejes de incertidumbre más influyentes en el modelo de población eran dependientes del procedimiento de ordenación.

KEYWORDS

GLM, Bluefin tuna, MSE, management procedures, axes of uncertainty

[1] Fisheries & Oceans Canada, Biological Station, 531 Brandy Cove Road, St. Andrews, NB E5B 2L9 CANADA. Email address of lead author: alex.hanke@dfo-mpo.gc.ca.
[2] AZTI, Marine Research, Basque Research and Technology Alliance (BRTA). Herrera Kaia, Portualdea z/g, 20110 Pasaia – Gipuzkoa

1. Introduction

At the May 2020 meeting of the Bluefin tuna Species Group, a general methodology for evaluating what operating model axes "matter" was discussed (Anon 2020, Section 7.6) and included using statistical methods which have an established model selection and validation process, helpful methods for effect visualization and can also help to determine whether there are significant interactions that render interpretation of the main effect of axes incomplete or misleading.

The reference grid of operating models provided in the R package ABTMSE (Carruthers 2020), on which cMPs were to be tested, reflects differences in Bluefin tuna population dynamics related to maturity at age, natural mortality at age, recruitment and mixing as well as uncertainty related to the scale of the spawning stock biomass and the goodness of fit to the length composition data. The fully cross-classified design matrix yielded 96 distinctly different operating models.

The objective here was to determine if a GLM approach could inform on the features (axes of uncertainty) of the OMs that were most influential. This information would then inform on decisions to reduce the number of OMs in the grid. It is also demonstrated how the method can provide a rank of cMP performance across the features of the OMs.

2. Methods

Data

The evaluation of the influence of the axes of uncertainty on cMP performance was based on models fit to the Br30 and AvC30 performance metrics. Br30 refers to the biomass ratio for a stock after 30 years of management using a given procedure whereas AvC30 is the average yield by the procedure. Given that there are two stocks (East and West), there are separate estimates of Br30 and AvC30 for each. In the case of AvC30 it was necessary to scale the values for each area to a common mean. A total of 9 cMPs were evaluated against the full 96 operating models and each yielded a data object with 192 rows. In order to reduce computation time, the deterministic version of the OMs with no observation error in the projected indices was used. The recruitment was also deterministic and the observation model had no error.

Candidate Management Procedures

A variety of empirical based procedures were tested to evaluate the influence of the axes of uncertainty more generally. The eight management procedures tested can be characterized as follows:

Regime: a constant TAC is set at 1,750,000 kg for the west area and 23,000,000 kg for the east area.

Rose23: the east area TAC is updated based on the current value of the combined MED_LAR_SUV and CAN_SWNS indices relative to a reference value and the current slope of the JPN_LL_NEAtl2 index. Additionally, a change point algorithm is also used to determine if the trend and state of the JPN_LL_NEAtl2 index is uncharacteristic of past values and adjusts TAC accordingly. The west area TAC is updated in a similar way using the combined GOM_LAR_SUV and CAN_SWNS indices for state and the JPN_LL_West2 index for trend and change point detection.

RoseEW: the TAC for both areas is updated using the same approach as for *Rose23* but the determination of state only depends on the GOM_LAR_SUV in the west area and the MED_LAR_SUV index for the east area. The other aspects are the same.

RoseEWb: similar to the previous MP but with a different index for trend and change point detection. Adaptive harvest cMPs RoseE and RoseW with JPLL used to detect regime shift

Fzero1EW: the TAC for the west area is determined from an estimate of F0.1 and SSB. The yield per recruit and estimate of F0.1 is calculated using west area estimates of weight and mortality at age used in the 2017 VPA. The selectivity at age is derived from the US_RR_66_114, US_RR_115_144 and US_RR_177 indices. The SSB estimate is based on the current value of the GOM_LAR_SUV index divided by an estimate of q. The east area procedure is similar using east area VPA inputs for the weight and mortality at age and the FR_AER_SUV2, US_RR_115_144 and MED_LAR_SUV indices to determine selectivity at age. The current MED_LAR_SUV index value divided by an estimate of q provides an estimate of SSB.

Fzero1EWb: the TAC is determined as above but with a change in the range of years used in the estimate of SSB and a different q.

FIEW: the updated TAC is the product of a constant F of 0.5 times an estimate of SSB adjusted by current status change. For the east area the SSB is the current MED_LAR_SUV index value divided by an estimate of q and the status change is the ratio of the current value of the MED_LAR_SUV index to its average over the most recent 3 years. A similar procedure is applied for the west area using the GOM_LAR_SUV index.

Fzero1RpSEW: this procedure is very similar to *Fzero1EW* with the addition of a scalar on q. The scalar increases q when the slope of the recruits: spawner ratio is greater than or equal to 0 and decreases it when it is below 0. For the east area the ratio is determined by a moving average of the FR_AER_SUV2 : MED_LAR_SUV indices and in the west area it is the US_RR_66_114 : GOM_LAR_SUV indices.

EUcMPmedianEW: this procedure uses 4 indices for each area to make the TAC adjustments. In the east area the indices are FR_AER_SUV2, MED_LAR_SUV, MOR_POR_TRAP and JPN_LL_NEAtl2 while for the west area they are GOM_LAR_SUV, US_RR_66_114, US_GOM_PLL2 and JPN_LL_West2. For each of the 4 indices in an area, the ratio of each of the 3 most recent years relative to the mean for the 2013:2016 period is calculated. The median of these 12 ratio values relative to a Target of 1 (west) or 0.75 (east) is used to adjust the TAC up or down but is limited to 20%.

EUcMPXXXEW: not available at time of publication.

Statistical Approach

The reference grid of operating models, on which the cMPs were tested, reflects differences in Bluefin tuna population dynamics related to maturity at age, natural mortality at age, recruitment and mixing as well as uncertainty related to the scale of the spawning stock biomass and the goodness of fit to the length composition data. These features that differ between OMs are treated as the independent variables in a GLM describing the variability in the dependent variables (Br30, AvC30). The factors and their levels are described as follows:

A) Recruitment (Regime)
1: West: h=0.6 to h=0.9 1975+, East: h=0.98 for 1987- to h=0.98 1988+
2: West: B-H h=0.6 all years, East: B-H h=0.7 all years
3: West: post 75+ changes to pre '75 after 10 yrs, East: 88+ to '50-87 after 10 yrs

B) Scale
--: mean SSB 15kt West, 200kt East
-+: mean SSB 15kt West, 400kt East
+-: mean SSB 50kt West, 200kt East
++: mean SSB 50kt West, 400kt East

C) Spawning Fraction/Natural Mortality (MatMort)
A: Younger spawning, High M
B: Older spawning, Low M

D) Mixing
I: Low West Stock Migration
II: High West Stock Migration

E) Length Composition (LenCom)
L: Low length composition weight of 1/20
H: High length composition weight of 1

F) Stock
1: East
2: West

Log-linear GLMs with Gaussian errors were fit to the dependent variables in a stepwise forward selection process (stepAIC, Venables and Ripley 2002). The complexity of the model space was bounded by a null model (y~1) as potentially the simplest explaining the variability in the data and a fully specified model with all two-way and three-way interactions (y~.^3). At each step the algorithm added the factor or interaction which reduced the AIC by the greatest amount. The selection process was halted when terms outside the model could not reduce the AIC by more than 2 units. A difference of 4 in AIC corresponds to ~7.4 times stronger evidence for the model with the lower AIC, while a difference in 2 corresponds to ~2.7 times stronger evidence.

This process was applied to the performance metric generated by each cMP and to the combined data. In the latter case, cMP was considered as an explanatory variable.

3. Results

Br30 Model Selection

Table 1 shows the outcomes of the model selection process for each of the 9 cMPs. Most of the models do not include Mixing as an influential factor but two did. Stock and MatMort were two of the other factors that were not included in at least one of the models. None of the models involved only main effects terms. Three of the models required two-way interactions to account for the variability in the response while the remainder also required three-way interactions. The inclusion of interactions implies that the single variables should not be interpreted as an "overall effect" on the dependent variable; i.e. the effect or influence of the single variable is conditional on the modifying effect of the factor it is interacting with.

AvC30 Model Selection

The model selection process applied to the AvC30 data is presented in Table 2 and demonstrates that the variability is accounted for by far simpler models with fewer main effects than for Br30. Regime, MatMort typically account for the most variability in the average catch with some models including Scale. Stock and Mixing were not selected in any of the models. No three-way interactions are important and relatively few two-way interactions making the interpretation of main effects conditional on fewer additional factors.

cMP Performance

The model selection process applied to the combined data from all the cMP data yielded fairly complex best fitting models.

The final models for Br30 and AvC30 were as follows:

Br30 ~ Regime + MP + Stock + Scale + LengthComp + SpawnMort + Mixing + Regime:MP + Stock:Scale + MP:Scale + MP:Stock + Stock:LengthComp + MP:LengthComp + Regime:LengthComp + Regime:Stock + MP:SpawnMort + Regime:Scale + Scale:LengthComp + Scale:SpawnMort + Regime:SpawnMort + Stock:Mixing + Stock:SpawnMort + Scale:Mixing + SpawnMort:Mixing + LengthComp:Mixing + LengthComp:SpawnMort + Regime:Mixing + MP:Stock:Scale + MP:Stock:LengthComp + Regime:Stock:LengthComp + Regime:MP:LengthComp + Regime:MP:Scale + Regime:Stock:Scale + Stock:Scale:LengthComp + MP:Scale:SpawnMort + MP:Scale:LengthComp + Regime:MP:SpawnMort + Regime:MP:Stock + Stock:Scale:SpawnMort + MP:Stock:SpawnMort + Regime:Scale:SpawnMort + Regime:Stock:SpawnMort + Stock:SpawnMort:Mixing + Regime:Scale:LengthComp + Scale:LengthComp:Mixing + Scale:LengthComp:SpawnMort + MP:LengthComp:SpawnMort + Scale:SpawnMort:Mixing + Stock:Scale:Mixing + Regime:Scale:Mixing + Regime:Stock:Mixing

AvC30 ~ Stock + MP + Regime + SpawnMort + Scale + LengthComp + Mixing + MP:Regime + MP:SpawnMort + Regime:SpawnMort + MP:Scale + Regime:Scale + SpawnMort:Scale + MP:LengthComp + Scale:LengthComp + Stock:Scale + Regime:LengthComp + Stock:MP + Stock:LengthComp + Scale:Mixing + MP:Mixing + SpawnMort:Mixing + MP:SpawnMort:Scale + MP:Scale:LengthComp + MP:Regime:SpawnMort + MP:Regime:Scale + MP:Regime:LengthComp + Regime:Scale:LengthComp + MP:Scale:Mixing + Regime:SpawnMort:Scale + MP:SpawnMort:Mixing

A comparison of the cMPs by Stock is provided in **Figure 1.** Note that it was necessary to estimate the effect by determining the unweighted averages over the non-focal factor levels including all higher order relatives of Stock and MP in the models. These plots show the tradeoffs in catch when maximizing the stock state and demonstrate the potential of this approach to rank cMP performance across the reference grid of OMs.

4. Conclusions

1. This exercise demonstrated that "*the axes of uncertainty that matter*" are cMP dependent. Consequently, changing or modifying the cMPs used in this exercise would change the determination of what matters. Furthermore, dropping an axis of uncertainty prior to selecting the best cMP for providing advice could affect which cMP is selected and indicates that it might be premature to simplify the axes too early in the cMP development and testing process.

2. Br30 models were much more complex than AvC30 models. Each of the different population model features (axes of uncertainty) were consequential for at least one of the Br30 models being tested, whereas Stock and Mixing did not account for significant variability in any of the AvC30 models. The determination of what matters is therefore dependent on the performance metric or metrics being evaluated. Although these are probably the most important performance metrics to focus on, the analysis suggested that it is important to agree on this before getting rid of any axis of uncertainty.

3. Performance related to stability and safety were not considered here and could be involved in the determination of what matters.

4. Many of the models included two and three-way interactions indicating that the effect of an axis of uncertainty is conditional on other axes. These dependencies further limit changes to the axes defining the OM grid.

5. While it was demonstrated that the GLM approach could be used to quantify cMP performance across the OM grid by averaging over all other factors and their interactions in the model, it is also possible to determine the effect on cMP ranking of fixing an axis at one of its factor levels. This exercise (not done) would demonstrate if cMP ranking is preserved when an axis of uncertainty is fixed in the population model.

6. Fairly simple cMPs required a complex GLM to account for the variability in the performance metric (compare the Br30 model of the constant TAC cMP *Regime* versus *Fzero1EW*). Consequently, if one is choosing among sufficiently sophisticated cMPs, the potential to drop axes can be greater.

References

Anon. 2020. REPORT OF THE 2020 ICCAT INTERSESSIONAL MEETING OF THE BLUEFIN TUNA SPECIES GROUP. Online, 14-22 May 2020.

Carruthers, T. (2020). ABTMSE: Atlantic Bluefin Tuna Management Strategy Evaluation. R package version 6.6.12.

Venables, W. N. & Ripley, B. D. (2002) Modern Applied Statistics with S. Fourth Edition. Springer, New York. ISBN 0-387-95457-0

Table 1. Br30 model selection results for 9 cMPs applied to the reference grid of 96 deterministic OMs. A forward selection stepAIC method with a threshold of 2 AIC units was used to choose the most influential factors and interactions up to 3rd order. Blue indicates significant main effect; grey indicates significant 2nd order interactions and orange indicates the 3rd order interactions with number indicating combinations of three.

MP: Regime

	Regime	Stock	LenCom	Scale	MatMort	Mixing
Regime	1,2,5					
Stock	1,2	3				
LenCom	2	3	4			
Scale	1,5	3	4			
MatMort	5		4			
Mixing						

MP: Rose23

	Regime	Stock	LenCom	Scale	MatMort	Mixing
Regime	1					
Stock	1					
LenCom						
Scale	1					
MatMort						
Mixing						

MP: RoseEW

	Regime	Stock	LenCom	Scale	MatMort	Mixing
Regime						
Stock						
LenCom						
Scale						
MatMort						
Mixing						

MP: RoseEWb

	Regime	Stock	LenCom	Scale	MatMort	Mixing
Regime						
Stock						
LenCom						
Scale						
MatMort						
Mixing						

MP: Fzero1EW

	Regime	Stock	LenCom	Scale	MatMort	Mixing
Regime	2,4,5,7					
Stock	2,4,5	1,3				
LenCom	2	1	6			
Scale	5,7	1,3	6			
MatMort	4,7	3	6			
Mixing						

MP: Fzero1EWb

	Regime	Stock	LenCom	Scale	MatMort	Mixing
Regime						
Stock						
LenCom						
Scale						
MatMort						
Mixing						

MP: FIEW

	Regime	Stock	LenCom	Scale	MatMort	Mixing
Regime	1,4,5					
Stock	1,4,5	2,3				
LenCom	1	2				
Scale	4	2,3				
MatMort	5	3				
Mixing						

MP: Fzero1RpSEW

	Regime	Stock	LenCom	Scale	MatMort	Mixing
Regime	2,4					
Stock	2,4	1,3				
LenCom	2	1				
Scale	4	1,3				
MatMort		3				
Mixing						

MP: EUcMPmedianEW

	Regime	Stock	LenCom	Scale	MatMort	Mixing
Regime						
Stock						
LenCom						
Scale						
MatMort						
Mixing						

MP: EUcMPXXXEW

	Regime	Stock	LenCom	Scale	MatMort	Mixing
Regime						
Stock						
LenCom						
Scale						
MatMort						
Mixing						

Table 2. AvC30 model selection results for 9 cMPs applied to the reference grid of 96 deterministic OMs. A forward selection stepAIC method with a threshold of 2 AIC units was used to choose the most influential factors and interactions up to 3rd order. Blue indicates significant main effect; grey indicates significant 2nd order interactions and orange indicates the 3rd order interactions with number indicating combinations of three. Note given the large differences in TAC for the east and west area, the catch data for each stock was centered on a mean of 1000 prior to model fitting.

MP: Regime

	Regime	Stock	LenCom	Scale	MatMort	Mixing
Regime						
Stock						
LenCom						
Scale						
MatMort						
Mixing						

MP: Rose23

	Regime	Stock	LenCom	Scale	MatMort	Mixing
Regime						
Stock						
LenCom						
Scale						
MatMort						
Mixing						

MP: RoseEW

	Regime	Stock	LenCom	Scale	MatMort	Mixing
Regime						
Stock						
LenCom						
Scale						
MatMort						
Mixing						

MP: RoseEWb

	Regime	Stock	LenCom	Scale	MatMort	Mixing
Regime						
Stock						
LenCom						
Scale						
MatMort						
Mixing						

MP: Fzero1EW

	Regime	Stock	LenCom	Scale	MatMort	Mixing
Regime						
Stock						
LenCom						
Scale						
MatMort						
Mixing						

MP: Fzero1EWb

	Regime	Stock	LenCom	Scale	MatMort	Mixing
Regime						
Stock						
LenCom						
Scale						
MatMort						
Mixing						

MP: FIEW

	Regime	Stock	LenCom	Scale	MatMort	Mixing
Regime						
Stock						
LenCom						
Scale						
MatMort						
Mixing						

MP: Fzero1RpSEW

	Regime	Stock	LenCom	Scale	MatMort	Mixing
Regime						
Stock						
LenCom						
Scale						
MatMort						
Mixing						

MP: EUcMPmedianEW

	Regime	Stock	LenCom	Scale	MatMort	Mixing
Regime						
Stock						
LenCom						
Scale						
MatMort						
Mixing						

MP: EUcMPXXXEW

	Regime	Stock	LenCom	Scale	MatMort	Mixing
Regime						
Stock						
LenCom						
Scale						
MatMort						
Mixing						

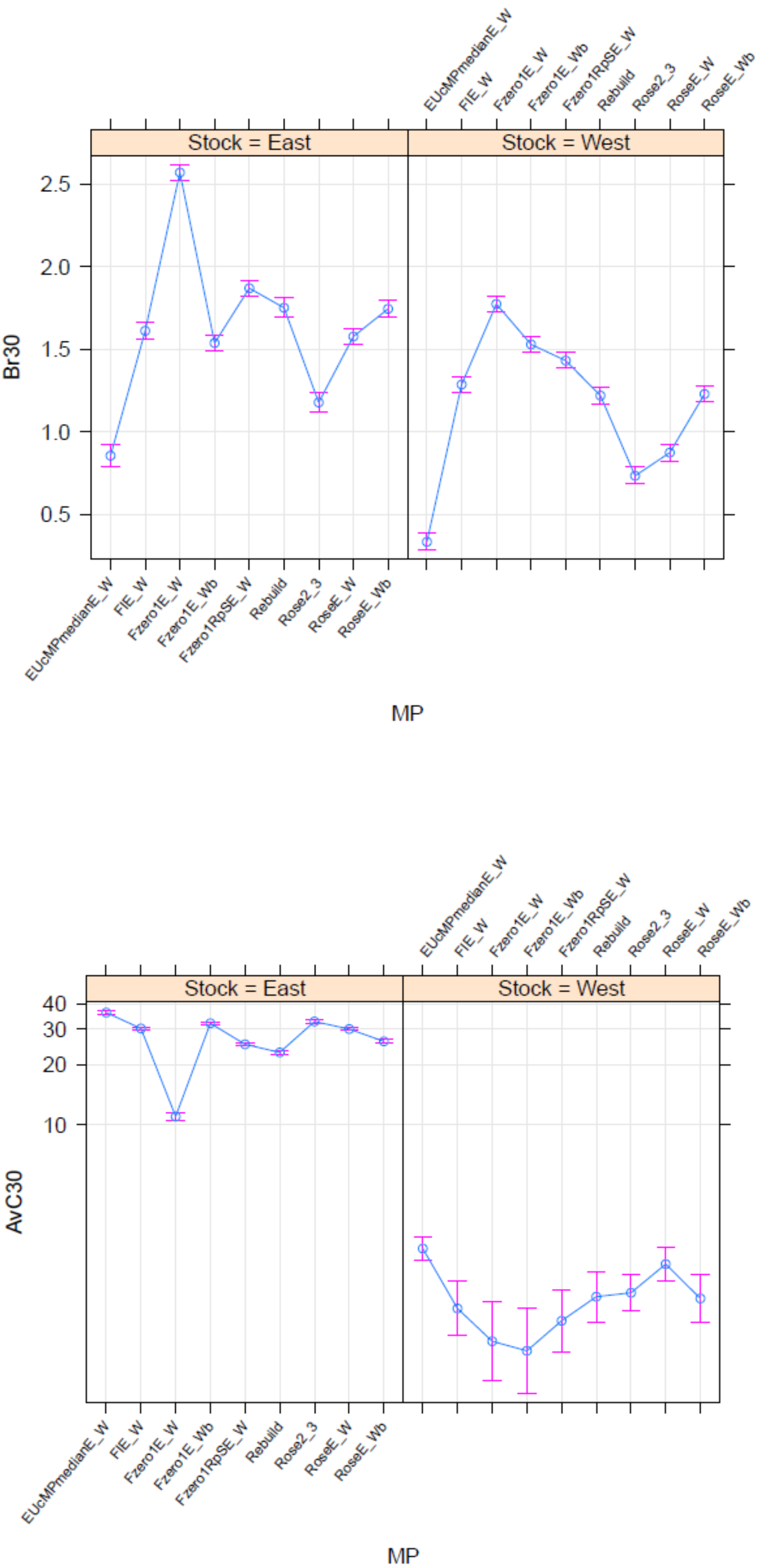

Figure 1. Stock and cMP effects plot for models fit to Br30 and AvC30. Note that the terms that have higher-order relatives in the model, the expected values is calculated based on averaging over those terms. Also AvC30 is plotted on the log scale.