SCRS/2022/104 Collect. Vol. Sci. Pap. ICCAT, 79(5): 165-177 (2022)

ECOTEST, A PROOF OF CONCEPT FOR EVALUATING ECOLOGICAL INDICATORS IN MULTISPECIES FISHERIES, WITH THE ATLANTIC LONGLINE FISHERY CASE STUDY

Quang C. Huynh[1], Tom Carruthers[2], Nathan G. Taylor[3]

SUMMARY

There is a need for rigorous science to inform decision makers for Ecosystem Based Fisheries Management (EBFM). It is important to establish challenging and plausible scenarios for ecosystem dynamics and then test whether current and potential indicators can reflect stock status. Without the validation of indicators and the testing of relevant policy guidance to mitigate ecosystem impacts, there is a credibility gap between scientific practitioners of ecosystem science and decision makers that need to defend their actions in large multi-party negotiations. A multi-species framework that supports tactical decision making can make significant progress towards the essential goals of EBFM. We present a management strategy evaluation framework called "EcoTest". This is an extension to openMSE software, used for single-species modeling, that simulates multi-species fisheries dynamics. A range of features are possible in EcoTest, such as the ability to evaluate current indicators as well as design new indicators and identify the conditions under which indicators operate reliably. Here we demonstrate the use of EcoTest using the Atlantic longline fishery as a case study.

RÉSUMÉ

Il est nécessaire de disposer d'une science rigoureuse pour informer les décideurs de la gestion des pêches fondée sur l'écosystème (EBFM). Il est important d'établir des scénarios stimulants et plausibles pour la dynamique des écosystèmes, puis de vérifier si les indicateurs actuels et potentiels peuvent refléter l'état des stocks. Sans la validation des indicateurs et le test des orientations politiques pertinentes pour atténuer les impacts sur les écosystèmes, il existe un manque de crédibilité entre les praticiens de la science des écosystèmes et les décideurs qui doivent défendre leurs actions dans de grandes négociations multipartites. Un cadre multi-espèces qui soutient la prise de décision tactique peut faire des progrès significatifs vers les objectifs essentiels de l'EBFM. Un cadre d'évaluation des stratégies de gestion appelé « EcoTest » est présenté dans ce document. Il s'agit d'une extension du logiciel openMSE, utilisé pour la modélisation mono-espèce, qui simule la dynamique des pêches multi-espèces. Une série de fonctionnalités sont possibles dans EcoTest, comme la possibilité d'évaluer les indicateurs actuels ainsi que de concevoir de nouveaux indicateurs et d'identifier les conditions dans lesquelles les indicateurs fonctionnent de manière fiable. Nous démontrons ici l'utilisation d'EcoTest en utilisant la pêcherie palangrière de l'Atlantique comme étude de cas.

RESUMEN

Se necesita una ciencia rigurosa para informar a los que toman las decisiones de la ordenación pesquera basada en el ecosistema (EBFM). Es importante establecer escenarios difíciles y plausibles para la dinámica del ecosistema y luego probar si los indicadores actuales y potenciales pueden reflejar el estado de los stocks. Sin la validación de los indicadores y la comprobación de las orientaciones normativas pertinentes para mitigar los impactos en los ecosistemas, existe una brecha de credibilidad entre los profesionales de la ciencia de los ecosistemas y los responsables de la toma de decisiones que deben defender sus acciones en las grandes negociaciones multipartitas. Un marco multiespecífico que apoye la toma de decisiones tácticas puede suponer un avance significativo hacia los objetivos esenciales de la EBFM. Presentamos un marco de evaluación de estrategias de ordenación denominado "EcoTest". Se

[1] Blue Matter Science Ltd, Vancouver, British Columbia, Canada. Email: quang@bluematterscience.com
[2] Blue Matter Science Ltd, Vancouver, British Columbia, Canada.
[3] ICCAT Secretariat. Corazón de Maria 8, Madrid, Spain.

trata de una extensión del software openMSE, utilizado para la modelación para una sola especie, que simula la dinámica de las pesquerías multiespecíficas. En EcoTest es posible realizar una serie de funciones, como la capacidad de evaluar los indicadores actuales, así como diseñar nuevos indicadores e identificar las condiciones en las que los indicadores funcionan de forma fiable. Aquí demostramos el uso de EcoTest utilizando la pesquería de palangre del Atlántico como caso de estudio.

KEYWORDS

Ecosystem based fisheries management, management strategy evaluation, harvest strategy

Introduction

Operationalizing initiatives related to EBFM has proven challenging for most t-RFMOs. ICCAT and other t-RMFOs rely on a system of indicators for EBFM. ICCAT compiles these in an Ecosystem Report Card, but these indicators have not frequently been validated to ensure that they are representative and sufficiently responsive to changes in the underlying species. For example, apparent decreases bycatch rates may be caused by either declines in the underlying bycatch population size, or be driven by changes in fleet distribution in search of target species leading some to conclude that CPUE generally cannot provide information needed to assess and manage communities or ecosystems (Maunder *et al.* 2006; Suzuki *et al.*, 1977; Bigelow *et al.* 2003).

Single-species stock assessments do not account for the impact of proposed management strategies for target species on associated bycatch species. These species may include for example, non-commercial species such as birds and turtles, and commercial species that are incidentally caught by some sectors of the fishery yet targeted in others, such as sharks and billfish. Therefore, a multi-species framework that supports tactical decision making can make significant progress towards the essential goals of EBFM. Such a framework can also support current efforts to develop indicators supporting EBFM. Simulation testing of indicators and management strategies linked to those indicators can validate their ability to characterize ecosystem impacts.

For simulation testing of indicators, it is important to establish challenging and plausible scenarios for ecosystem dynamics. In this way, it will be possible to test if current and yet-to-be-developed indicators can be expected to correctly detect ecosystem dynamics of interest. In situations where indicators detect potential problems, it is also necessary to test mitigation measures, i.e., management procedures (Punt *et al.* 2016), in order to meet conservation objectives. Management procedures (MPs) linked to the indicators so that a clear pathway from data to management policies has been evaluated. For example, in response to declining bycatch rates, the prescription of a size limit for a target species, or a time-area closure could be tested. Most indicators currently being considered for EBFM at ICCAT, along with harvest strategies that respond to changes to such indicators, have been validated. Without the validation of indicators and the testing of relevant policy guidance to mitigate ecosystem impacts, it is unclear how an indicator system can be used for decision making.

The complexity of multi-species modeling has been the biggest obstacle for simulation testing. Here, we present a management strategy evaluation framework: "EcoTest". It is an extension to openMSE software used to simulate multi-species fisheries dynamics. A range of features are possible in EcoTest, such as the ability to simulate the performance of current indicators as well as design new indicators. Bycatch indicators and policy options can also be formally and rigorously evaluated in EcoTest.

Here we describe the use of EcoTest using the Atlantic longline fishery as a preliminary case study example, testing how several reference MPs applied in primary (target) species affect the conservation and yield performance in the secondary (bycatch) species.

1. Methods

The EcoTest framework applied here consisted of four steps:

(1) Develop operating models representing a range of species caught in a multispecies fishery
(2) Develop hypotheses regarding the fishing dynamics and how catch of one species affects other species
(3) Develop projections, accounting for the fishing dynamics specified in Step 2
(4) Evaluate outcomes from simulation testing and whether there is proposed indicators can differentiate between good and poor outcomes

1.1 Operating models

We developed operating models (OMs) for six ICCAT species: bigeye tuna (BET), North Atlantic swordfish (SWO), blue shark (BSH), North Atlantic shortfin mako shark (SMA), white marlin (WHM), and blue marlin (BHM). Recent Stock Synthesis 3 assessments produced by the respective working groups were used to specify the operating model dynamics, (**Table 1**, **Figure 1**), including growth, fecundity, natural mortality, and estimated historical trajectories in abundance, recruitment, and fishing mortality. A two-fishing gear structure ("Longline" and "Other") was created in the operating models by aggregating fleets from the assessments. Fleets, typically stratified by country and spatial area, were identified by gear, either based on longline or other surface gear (purse seine, bait boat, etc.), and the operating model calculated the selectivity of the gear by aggregating the fishing mortality (F) of the constituent assessment fleets (**Figure 2**).

1.2 Fishing dynamics

The historical fishing mortality rates were used to develop a model to describe the relationship between longline fishing mortality among species. Multivariate linear regression was used to predict F for one subset of species based on the F from another subset. Here, the predictor species in the regression were BET and SWO. They were identified as "primary" species of high economic importance. Fishing mortality for these species could be a predictor of overall longline effort. The F for the other four species (BSH, SMA, WHM, BUM) were the response variables in the regression. They were identified as "secondary" species that may be incidentally caught or targeted less frequently than primary species. Fishing mortality of secondary species may respond to regulations and dynamics related to the primary species.

A regression predicted fishing mortality as a linear relationship:

$$\overrightarrow{F_y}^{s} = \overrightarrow{F_y}^{p}\boldsymbol{\beta}$$

where the fishing mortality for secondary species (a vector denoted by superscript "s") in year *y* is predicted from the value in the primary species (a vector denoted by superscript *p*), and the matrix of estimated slope coefficients $\boldsymbol{\beta}$.

1.3 Projections

For demonstration purposes, deterministic 50-year projections were run with three fishing scenarios to illustrate outcomes arising from the proposed regression model. First, the primary species were fished at F_{MSY}, and the longline F of the secondary species were predicted by the linear relationship estimated in the previous section (this scenario is labeled "F_{MSY}_primary"). Second, the primary species were fished at 50% F_{MSY}, with the secondary species longline F predicted by the linear relation ("0.5F_{MSY}_primary"). For both, the ratio of F between the Longline and Other gears in the primary species were held constant in the projection period to that in the most recent historical year. For the secondary species, the Other F was also held constant in the projection period to the value in the most recent year. To allow for comparison, the third scenario set F = 0 (no fishing mortality) for all stocks. All projections were run without process error, i.e., no variability in the recruitment predicted by the stock-recruitment relationship.

1.4 Indicators

From the projections, indicators can be calculated from simulated observations and operating model dynamics. Here we simulate two indicators: longline CPUE and longline mean length (of catches). The CPUE was proportional to the biomass vulnerable to the longline fishery, while the mean length was calculated from the simulated catch-at-length vector generated from a multinomial sampling distribution. To illustrate the effect of fishery interactions in EcoTest more clearly, no observation error was modeled in the CPUE and the sample size was high (10,000) in order to approximate the vulnerable length distribution in the operating model. In this way the relationship of the indicator to the underlying population is clear and not obscured by observation error.

2. Results

Historical fishing mortality in the longline fishery are positively correlated to some extent among the six species (**Figure 3**). Strong correlations occur for some species pairs, e.g., between SWO-BSH and between SMA-BET, whereas several others have weaker correlations with more noise in the relationship, e.g., BET-BSH. The relationship in some pairs appeared to be time-varying, which different slopes in the early time period compared to the recent period, e.g., between BUM and each primary species. This suggests that the fishery developed at different rates historically, with increasing bycatch exploitation rates experienced earlier for BUM.

The linear relationship showed different rates of change in F among secondary species (**Figure 4**). For example, the change in BSH F is strongly dependent on the F for SWO, but very little on BET F. On the other hand, the gradient in F of the other three secondary species is related to values from both BET and SWO. WHM had the steepest changes in F out of the four secondary species.

With the F_{MSY} fishing scenario for the primary species, three out of the four secondary species, i.e., BUM, SMA, and WHM, were in an overfished state ($SSB/SSB_{MSY} < 1$) at the end of the projection (**Figure 5**). For the operating model that we investigated, BUM and WHM were already overfished at the beginning of the projection and remained so, while SSB of SMA declined during the projection. In the 50% F_{MSY} fishing scenario, spawning biomass for all species were higher (**Figure 6**). For BUM, SMA, and WHM, the stocks were near SSB_{MSY} by the end of the projection. For both fishing scenarios, BSH remained above SSB_{MSY}.

Stock biomass increased the fastest in the no fishing scenario for five of the six species. For SMA, the stock declined during the first third of the projection regardless of future fishing mortality. This appears to be a result of juvenile mortality caused by low steepness and selectivity of predominantly immature animals (**Figure 2**). Recovery in the spawning biomass is delayed until the survival of new cohorts is increased.

The longline CPUE and mean length increased roughly linearly over time for three of the four species (except SMA; **Figure 7**). The rates of increase are inversely proportional to fishing mortality. Between the F_{MSY} and 50% F_{MSY} scenarios, the magnitude in the indicators is similar initially in the projection and more time is needed in order to produce contrast in the indicator space. For SMA, the indicators are out of phase with stock status owing to differences between selectivity and maturity. Spawning biomass declines during the projection because past fishing mortality on immature animals will reduce survival to maturity, and in turn, reduces recruitment in the middle of the projection. Accordingly, the CPUE and mean length decrease in the middle of the projection following an initial increase. This creates a hook-like trajectory during the projection in contrast to the linear increases observed in the other species (**Figure 7**).

3. Discussion

With projections of a multispecies fishery under various scenarios (for example any adopted MP in those fisheries), the corresponding behavior of indicators can be evaluated. Indicators such as mean length and ratios of indices have been developed using equilibrium assumptions, e.g., constant recruitment and mortality (Beverton and Holt 1956; Ricker 1975). The MSE approach provides the opportunity to simulation test indicators and their suitability as a proxy for population status. Equilibrium assumptions are violated where stocks may experience changes in mortality over time, e.g., through rebuilding plans or multispecies fishery interactions, and recognizing density-dependent recruitment. For example, either an increase in mortality or episodic recruitment can decrease mean length. Here, selectivity pattern and low stock-recruit steepness of SMA generates different transient behavior of the indicators compared to the other five stocks evaluated in the species complex. Evaluation in EcoTest identifies when these indicators and others, such as species catch ratios, are informative and sufficiently responsive to changes in status and exploitation; this can also show when indicators are not responsive to changes in status and exploitation.

This initial demonstration used recent assessments for generating operating models for billfish, tuna, and shark species. Operating models can also be developed for other species that are not assessed. In a data-limited situations, operating models need not make statements about current status (Sagarese *et al.* 2018). Rather, these operating models can be the basis for evaluating which indicators are likely to be valuable and robust under uncertain conditions. Overall, a testing platform that uses population dynamics models can increase confidence for use of indicators for assigning status.

3.1 Next steps

Modeling a multispecies fishery require specifying the relationship between effort and fishing mortality among species in the complex. Here, we used a "top-down approach" to look at broad trends in historical fishing mortality based on estimates from a limited number of single-species assessments. Overall, there was a positively correlated relationship for fishing mortality among species based on the expansion of the longline fishery during the 20th century, which was modeled broadly as a linear function in the projection. Further exploration can be used to specify more complex relationships, e.g., power function (Hilborn and Walters 1992) or stepwise functions in fishing mortality. By reducing the correlation in mortality, projections can explore how much the fishery needs to avoid incidental catch to avoid harm to incidental species.

Fisheries frequently switch targeting between species and effort may depend on availability and spatial distribution of the underlying stock. EcoTest has capabilities for developing spatial operating models to describe abundance hotspots and coldspots, and potential range contraction. Likewise, fishing mortality for a fleet that targets based on availability can be distributed among regions as:

$$F_{y,r} = E_y \frac{VB_{y,r}^{\lambda}}{\sum_r VB_{y,r}^{\lambda}}$$

where r indexes spatial area, E is an index of overall fishing effort, VB is the vulnerable biomass, and λ is the targeting parameter. These features were not included in the demonstration, but can be more representative of the underlying fishery dynamics. Spatial fishery data such as the Task 1 and Task 2 ICCAT statistical databases provide information on parameterizing these spatial operating models (**Figure 8**). This "bottom-up approach" models bycatch hotspots and coldspots (Mannocci *et al.* 2020; Mucientes *et al.* 2022), which are dynamics more complex than what is modeled in single species assessments. In turn, indicators used for EBFM can be simulated and evaluated to test their core assumptions. For example, history has shown the need for taking great care with analyzing fishery CPUE (Walters 2003). In a spatial operating model, we can test the extent to which CPUE is proportional to stock abundance, an emergent property of the spatial dynamics (stock movement and fishery targeting) incorporated in the operating models.

Finally, there has been a lack of coordination between management measures that respond to EBFM indicators (Juan-Jordá *et al.* 2018). Management actions linked to indicator values can be operationalized in MPs and tested in EcoTest as an extension of management strategy evaluation. Evaluation of these MPs against ecosystem objectives will provide the justification for EBFM implementation.

Acknowledgements

The authors acknowledge the Ocean Foundation for funding support.

References

Anonymous. 2017. Report of the 2017 ICCAT shortfin mako assessment meeting. SCRS/2017/007 Collect. Vol. Sci. Pap. ICCAT, 74(4): 1465-1561.

Anonymous. 2018. Report of the 2018 ICCAT blue marlin stock assessment meeting. SCRS/2018/008. Collect. Vol. Sci. Pap. ICCAT, 75(5): 813-888.

Anonymous. 2020. Report of the 2019 ICCAT white marlin stock assessment meeting. SCRS/2019/009. Collect. Vol. Sci. Pap. ICCAT, 76(4): 97-181.

Anonymous. 2021. Report of the 2021 bigeye stock assessment meeting. SCRS/2021/011. Collect. Vol. Sci. Pap. ICCAT, 78(2): 335-485.

Beverton, R. J. H., and S. J. Holt. 1956. A review of methods for estimating mortality rates in fish populations, with special reference to sources of bias in catch sampling. Rapports et Procès-Verbaux des Réunions Conseil International pour l'Exploration de la Mer 140:67–83.

Bigelow K., Maunder M.N., and Hinton, M. 2003. Comparison of deterministic and statistical habitat-based models to estimate effective longline effort and standardized CPUE for bigeye and yellowfin tuna. 16th Meeting of the Standing Committee on Tuna and Billfish, RG-3. 18 pp.

Courtney, D. 2016. Preliminary Stock Synthesis (Ss3) Model Runs conducted for North Atlantic blue shark. SCRS/2015/151. Collect. Vol. Sci. Pap. ICCAT, 72(5): 1186-1232.

Courtney, D., Cortes, E., and Zhang, X. 2017. Stock Synthesis (Ss3) Model Runs conducted for North Atlantic shortfin mako shark. SCRS/2017/125 Collect. Vol. Sci. Pap. ICCAT, 74(4): 1759-1821.

Hilborn, R., and Walters, C.J. 1992. Quantitative fisheries stock assessment: choice, dynamics, and uncertainty. Chapman and Hall, New York.

Juan-Jordá, M.J., Murua, H., Arrizabalaga, H., Dulvy, N.K., and Restrepo, V. 2018. Report card on ecosystem-based fisheries management in tuna regional fisheries management organizations. Fish Fish. 19: 321-339.

Mannocci, L., Forget, F., Travassos Tolotti, M., Bach, P., Bez, N., Dagorn, L. 2020. Predicting bycatch hotspots in tropical tuna purse seine fisheries at the basin scale. Global Ecology and Conservation 24: e01393.

Maunder, M.N., Sibert, J.R., Fonteneau, A., Hampton, J., Kleiber, P., and Harley, S.J. 2006. Interpreting catch per unit effort data to assess the status of individual stocks and communities. ICES Journal of Marine Science 63: 1373-1385.

Mucientes, G., Vedor, M., Sims, D.W., and Queiroz, N. 2022. Unreported discards of internationally protected pelagic sharks in a global fishing hotspot are potentially large. Biological Conservation 269: 109534.

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.

Ricker, W. E. 1975. Computation and interpretation of biological statistics of fish populations. Fisheries Research Board of Canada Bulletin 191.

Sagarese, S.R., Rios, A.B., Cass-Calay, S.L., Cummings, N.J., Bryan, M.D., and Stevens, M.H. 2018. Working Towards a Framework for Stock Evaluations in Data-Limited Fisheries. North American Journal of Fisheries Management 38: 507-537.

Schirripa, M. 2018. Current status of the blue marlin (*makaira nigricans*) stock in the Atlantic Ocean 2018: Pre-Decisional Stock Assessment Model. SCRS/2018/097. Collect. Vol. Sci. Pap. ICCAT, 75(5): 1026-1050.

Schirripa, M. 2020. Current status of the white marlin (*kajikia albida*) stock in the Atlantic Ocean 2019: Predecisional Stock Assessement Model. SCRS/2019/110. Collect. Vol. Sci. Pap. ICCAT, 76(4): 265-292.

Schirripa, M.J. and Hordyk, A. 2020. Migrating the North Atlantic swordfish stock assessment model to an updated version of stock synthesis with analysis of the current minimum size regulation. SCRS/2020/159 Collect. Vol. Sci. Pap. ICCAT, 77(3): 654-668.

Suzuki Z., Warashina Y., and Kishida, M. 1977. The comparison of catches by regular and deep tuna longline gears in the western and central equatorial Pacific, Bulletin of the Far Seas Fisheries Research Laboratory 15: 51-89.

Walters, C. 2003. Folly and fantasy in the analysis of spatial catch rate data. Can. J. Fish. Aquat. Sci. 60: 1433–1436.

Table 1. Brief description of SS3 stock assessments used to develop operating models.

Stock	***Reference***	***Description***
BET	Anonymous (2021)	M = 0.2 and steepness = 0.8
SWO	Schirripa and Hordyk (2020)	Base Model (M = 0.2 and steepness = 0.75)
BSH	Courtney (2016)	Run 6 (Best convergence diagnostics, less weight to the length composition likelihood)
SMA	Anonymous (2017), Courtney *et al.* (2017)	Run 1, steepness = 0.354
WHM	Anonymous (2020), Schirripa (2020)	Model 6 (Use all CPUE indices except EU_Spain longline, without a catch multiplier, with variance reweighting)
BUM	Anonymous (2018), Schirripa (2018)	Base Model (M = 0.122 and steepness = 0.50)

Table 2. Estimated slope coefficients (β) for the linear relationship in historical longline F.

	Primary Species	
Secondary Species	***BET***	***SWO***
BSH	0.03	0.66
SMA	0.91	0.08
WHM	1.55	1.16
BUM	0.43	0.08

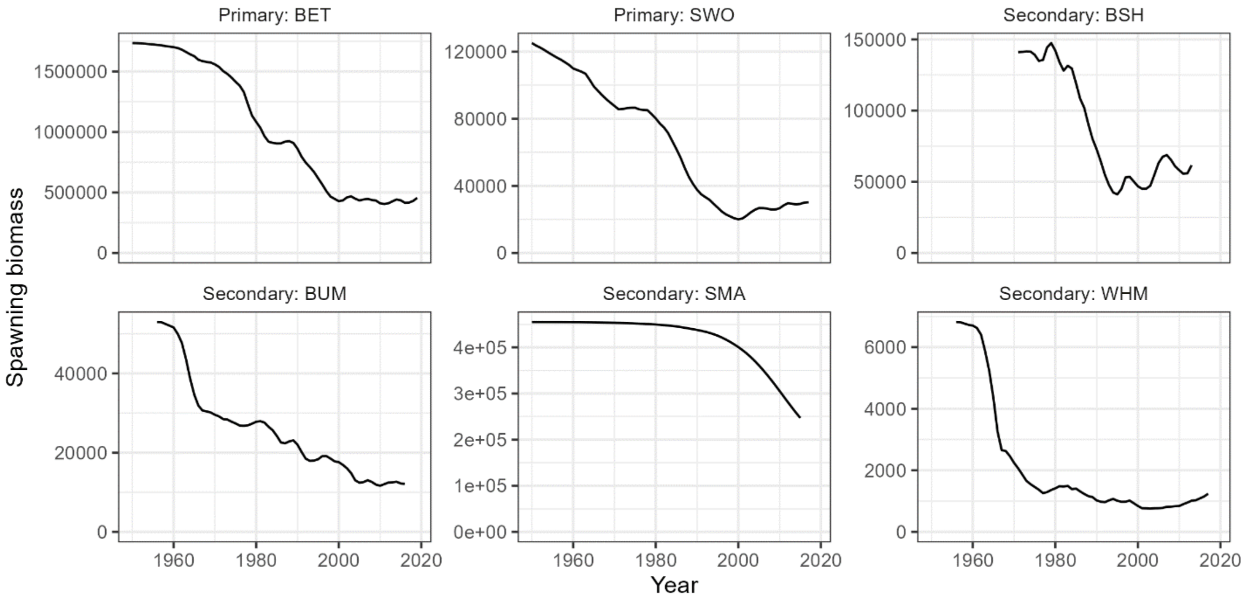

Figure 1. Spawning biomass estimated in SS3 assessments. See **Table 1** for description and references.

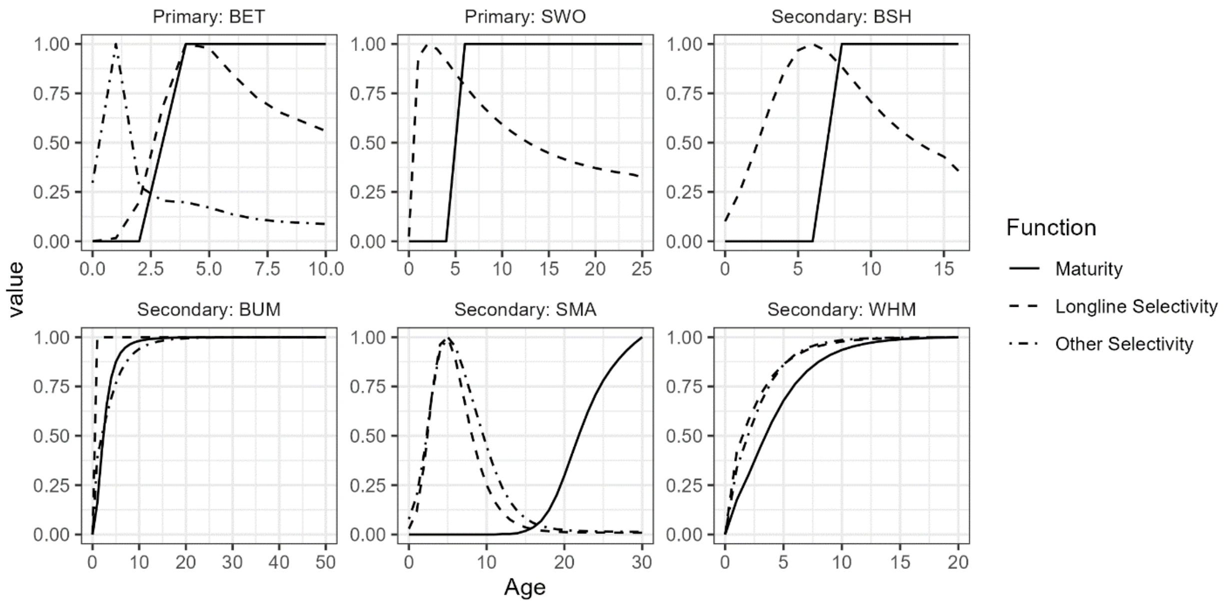

Figure 2. Aggregate selectivity (relative to maturity) of the longline and other (non-longline) gears from the operating models.

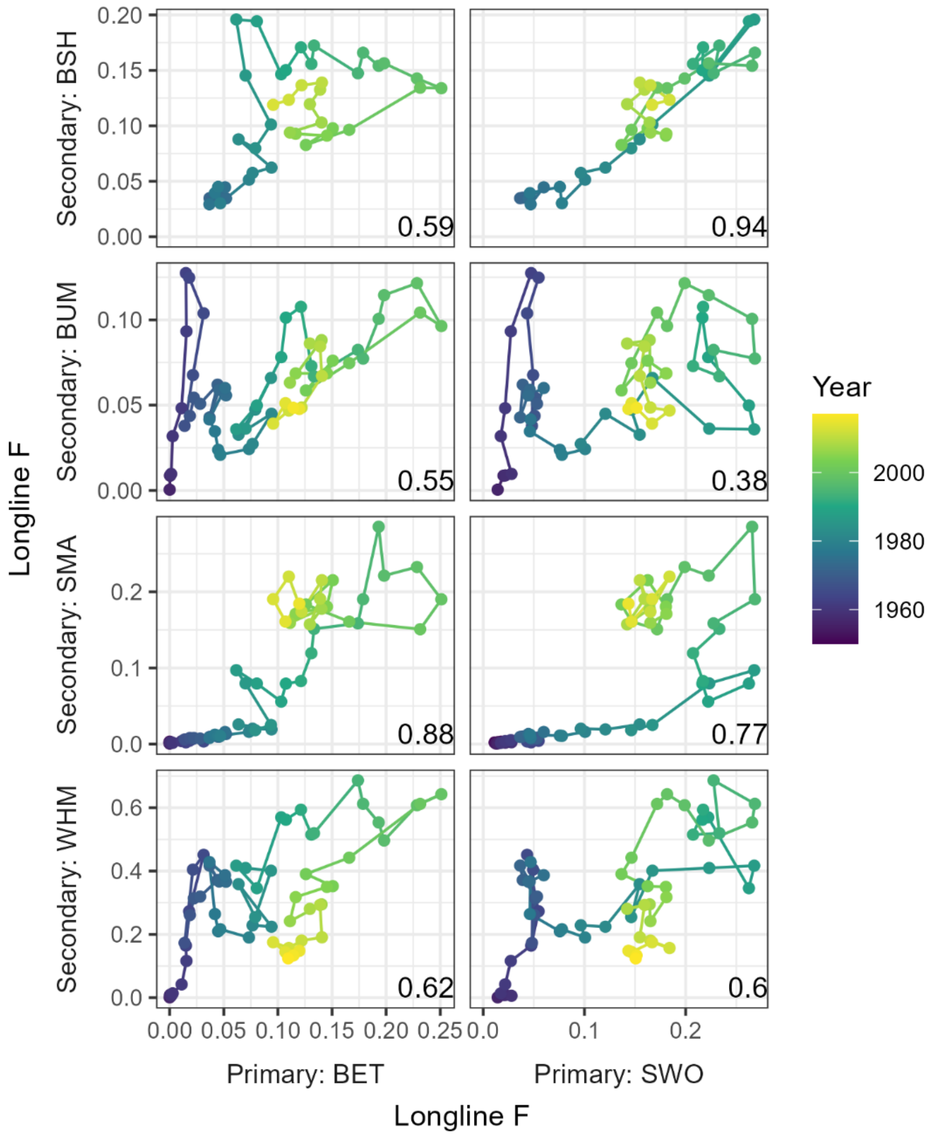

Figure 3. Phase plots over time of historical longline apical fishing mortality between the primary and secondary species. Numbers in lower right report the correlation in F in each panel.

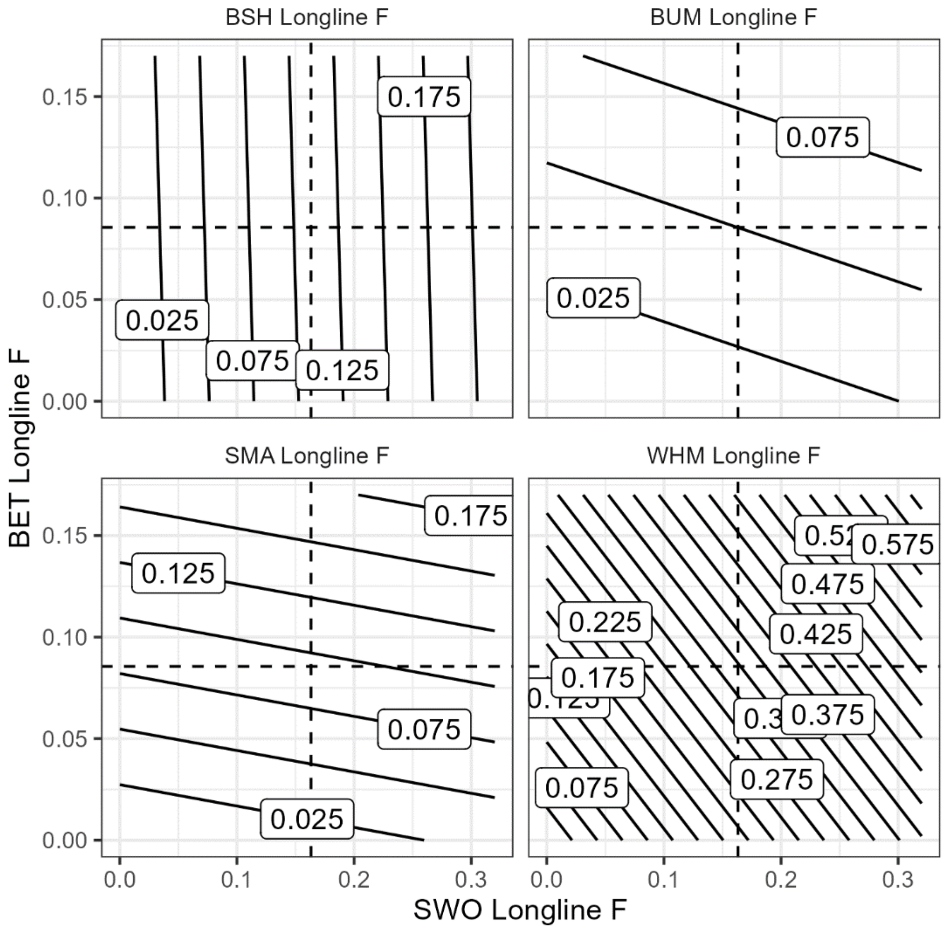

Figure 4. Contour plot of longline F predicted for the secondary species (in solid lines) by the F of the primary species using the linear relationship. Dotted horizontal and vertical lines denote FMSY for the primary species.

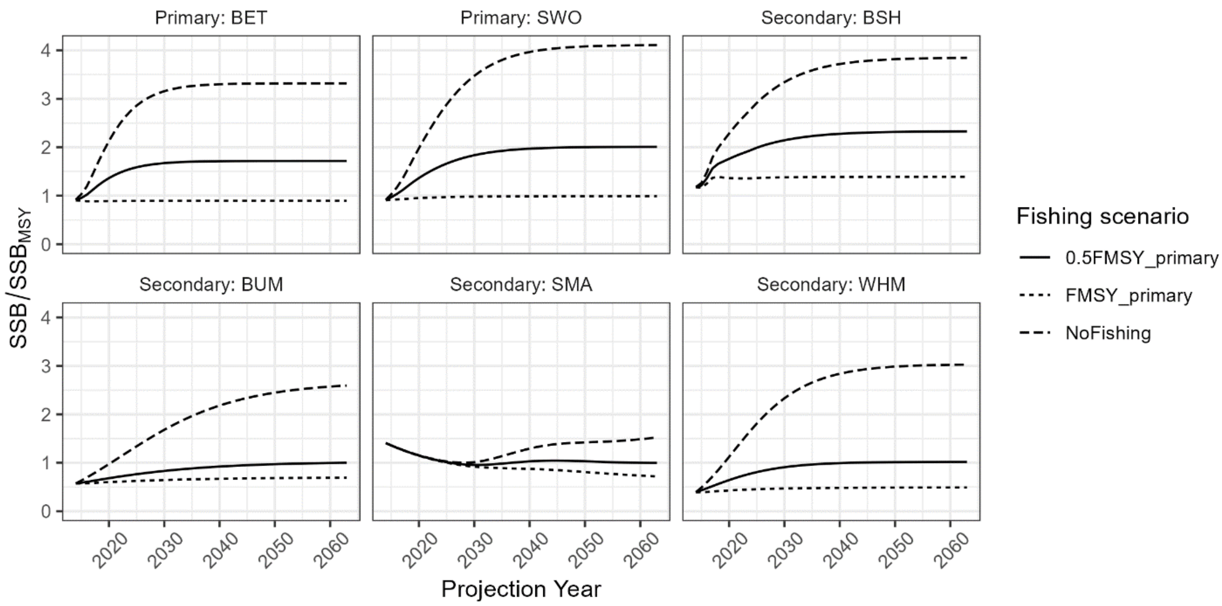

Figure 5. Projected SSB/SSB_{MSY} for the six stocks under the 3 fishing scenarios.

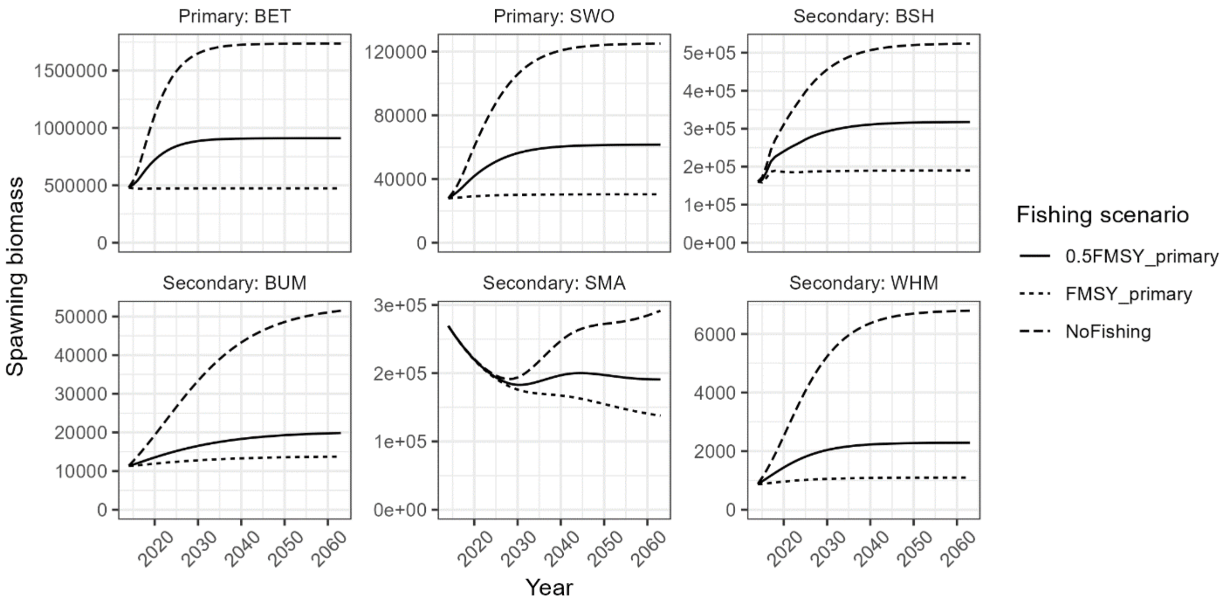

Figure 6. Projected spawning biomass for the six stocks under the 3 fishing scenarios.

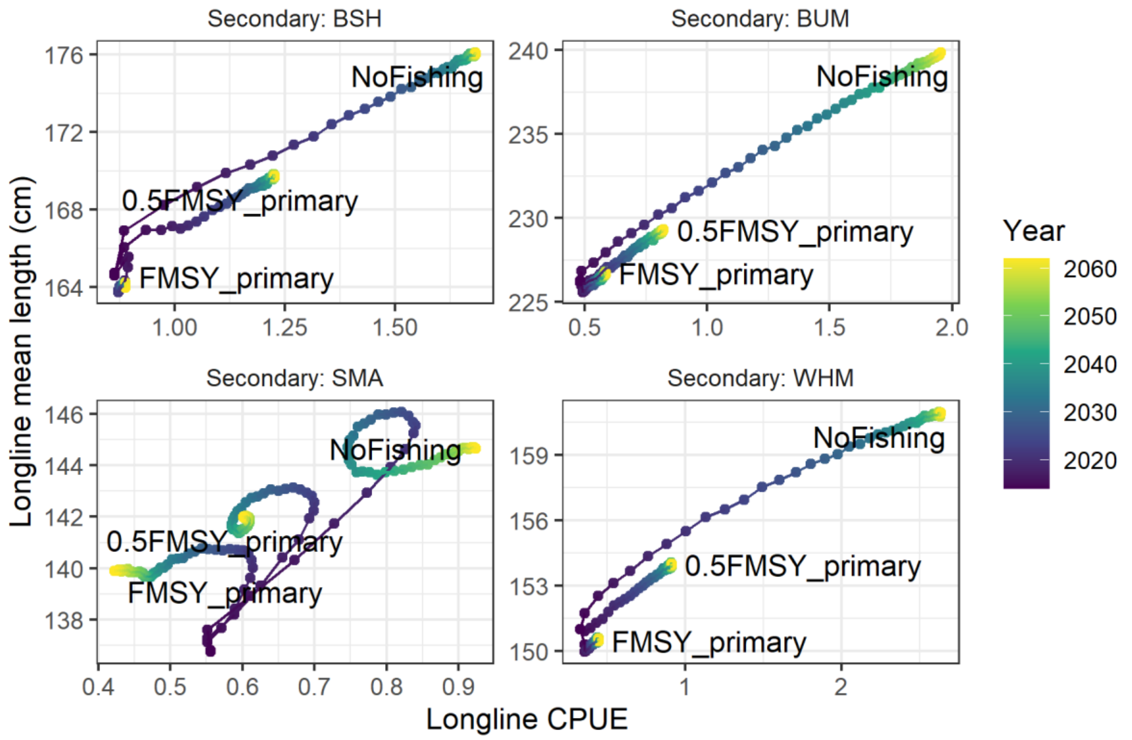

Figure 7. Phase plot of simulated longline mean length vs CPUE (relative by species) for the secondary species in the projections under the three fishing scenarios.

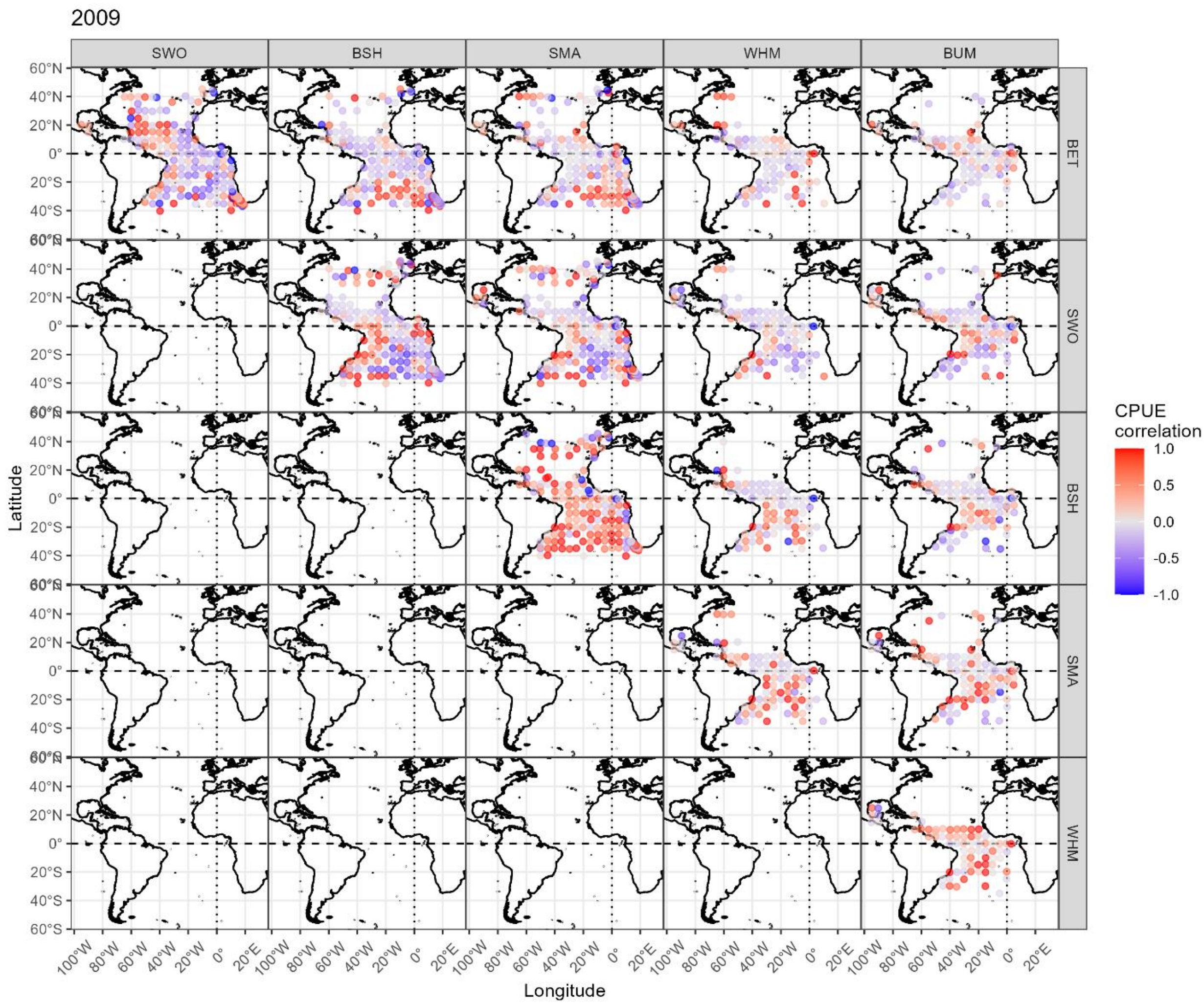

Figure 8. Spatial correlation in longline CPUE (kg/hook) between six species for 2009. CPUE were obtained from the ICCAT Task 2 CE database where records are aggregated to spatial grid (5˚ x 5˚), flag, and month.