Collect. Vol. Sci. Pap. ICCAT, 81(6), SCRS/2024/028: 1-12 (2024)

A REVIEW OF OBJECTIVES, REFERENCE POINTS, AND PERFORMANCE INDICATORS FOR MANAGEMENT STRATEGY EVALUATION AT TRFMOS

N.G. Taylor, S. Miller, and N. Duprey

SUMMARY

We reviewed the management measures related to management strategy evaluation processes at the International Commission for the Conservation of Atlantic Tunas, the Inter-American Tropical Tuna Commission, the Indian Ocean Tuna Commission, the Western and Central Pacific Fisheries Commission, and the Commission for the Conservation of Southern Bluefin Tuna. We defined a set of data fields to create a database of Performance Indicators and associated probability requirements, as well as objectives for desired stock status, yield, and safety (as expressed by limit reference points, LRPs), and variability in yield. We show that with respect to yield and status criteria, the tRFMOs have defined relatively consistent objectives in that they are striving to maximize catches and achieve maximum sustainable yield. While LRPs were not consistently defined among tRFMOs, the establishment of probabilities in avoiding them were relatively consistent. Finally, the criteria used to measure variability in yield and the magnitude of the variance permitted in management procedure (MP) design varied greatly across the tRFMOs.

RÉSUMÉ

Nous avons examiné les mesures de gestion liées aux processus d'évaluation de la stratégie de gestion (MSE) de la Commission internationale pour la conservation des thonidés de l'Atlantique (ICCAT), de la Commission interaméricaine du thon tropical (IATTC), de la Commission des thons de l'océan Indien (IOTC), de la Commission des pêches pour le Pacifique occidental et central (WCPC) et de la Commission pour la conservation du thon rouge du Sud (CCSBT). Nous avons défini un ensemble de champs de données pour créer une base de données d'indicateurs de performance et d'exigences de probabilité associées, ainsi que des objectifs pour l'état du stock, la production et la sécurité souhaités (tels qu'exprimés par les points de référence limites, LRP), et la variabilité de la production. Nous montrons qu'en ce qui concerne les critères de production et d'état, les ORGP thonières ont défini des objectifs relativement cohérents dans la mesure où elles s'efforcent de maximiser les captures et d'atteindre une production maximale équilibrée. Bien que les LRP n'aient pas été définis de manière cohérente par les ORGP thonières, l'établissement de probabilités pour les éviter a été relativement cohérent. Enfin, les critères utilisés pour mesurer la variabilité de la production et l'ampleur de la variance autorisée dans la conception des procédures de gestion (MP) varient considérablement d'une ORGP thonière à l'autre.

RESUMEN

Revisamos las medidas de ordenación relacionadas con los procesos de evaluación de estrategias de ordenación de la Comisión Internacional para la Conservación del Atún Atlántico (ICCAT), la Comisión Interamericana del Atún Tropical (IATC), la Comisión del Atún para el Océano Índico (IOTC), la Comisión de Pesca del Pacífico Occidental y Central (WCPFC) y la Comisión para la Conservación del Atún Rojo del Sur (CCSBT). Definimos un conjunto de campos de datos para crear una base de datos de indicadores de desempeño y requisitos de probabilidad asociados, así como objetivos del estado deseado del stock, de rendimiento y de seguridad (expresados mediante puntos de referencia límite, LPR) y de la variabilidad del rendimiento. Demostramos que, con respecto a los criterios de rendimiento y estado, las OROP de túnidos han definido objetivos relativamente coherentes en el sentido de que se esfuerzan por maximizar las capturas y alcanzar el rendimiento máximo sostenible. Aunque los LRP no se definieron de forma coherente entre las OROP de túnidos, el establecimiento de probabilidades para evitarlos fue relativamente coherente. Por último, los criterios utilizados para medir la variabilidad del rendimiento y la magnitud de la varianza que se permite en el diseño del procedimiento de ordenación (MP) variaron enormemente entre las OROP de túnidos.

KEYWORDS

Fishery management, Fishery policy, Fishery regulations, Quote regulations, Management Strategy Evaluation, Performance Indicators

1. Introduction

Regional fisheries management organizations dedicated to tunas (tRFMOs), including the Indian Ocean Tuna Commission (IOTC), the Inter-American Tropical Tuna Commission (IATTC), the Convention for the Conservation of Southern Bluefin Tuna (CCSBT), the Western and Central Pacific Fisheries Commission (WCPFC), and the International Commission for the Conservation of Atlantic Tunas (ICCAT), are pursuing management strategy evaluation (MSE) to develop management procedures (MPs) for many stocks under their jurisdiction. To pursue MSE, best practice calls for managers to first set specific and measurable management objectives at the onset (Punt *et al.* 2016). This includes defining reference points and performance indicators. With the aim of understanding the criteria used to define the performance of MPs in MSE, we summarize common practices for reference points and performance indicators that are considered at the tRFMOs worldwide.

To quantify each management objective, MSE processes require performance indicators (PIs). These determine performance criteria to evaluate the relative likelihood of success of candidate management procedures. They include measurable objectives, time frames, and probabilities. At ICCAT and other tRFMOs, the current practice is to explore unique PIs for each stock. If the plan is to do MSE for many fisheries, then efficiency becomes more important. Summarizing customary practice at tRFMOs might reveal if there is an emerging equifinality that could be considered as a reasonable basis for a common set of standards.

First it might help to explore what reference points are. As an example, ICCAT Rec. 15-07 defines limit reference points, targets, and thresholds as follows: 1. A limit is a conservation reference point based on a level of biomass (B_{LIM}) that should be avoided, considering that beyond such limits, the sustainability of the stock may be in danger. 2. A target is a management objective based on a level of biomass (B_{TARGET}) or a fishing mortality rate (F_{TARGET}) that should be achieved and maintained. And 3. A threshold is a level of biomass ($B_{THRESHOLD}$) reflecting the precautionary approach that triggers pre-agreed management actions to reduce the risk of breaching the limits. The recommendation further argues that thresholds should be set sufficiently far away from limits so that there is low probability that the limits could be exceeded.

To interpret ICCAT's Rec. 15-07 and measures like it, it is helpful to distinguish between the terminology of the Best Assessment (BA) paradigm (Parma 2002, Butterworth 2007) and Management Strategy Evaluation (MSE, Punt *et al.* 2016). We follow Cox *et al*. 2013 and use the terms operational control points: these are data values or model estimates that might provoke a change in the application of MP or in the BA paradigm. They are distinct from biological reference points that are quantities defined by Operating Models for MSE. This separates quantities that are objectives like LRPs and Target Reference Points from values that are empirical or quantities like $B_{THRESHOLD}$ at which management measures may change (i.e., an operational control point). This is important because i) not all MPs require that there be a threshold, for example most empirical MPs and fixed harvest rate strategies (Hall *et al.* 1988; Walters and Martell 2004; Carruthers *et al.* 2023); and ii) because the operational control points may themselves be estimated unreliably (Ludwig and Walters, 1984; NRC 1998, Magnusson and Hilborn 2007).

Distinguishing between operational control points and reference points avoids conflating objectives and strategies. In other words, B_{TARGET} may or may not be equal to $B_{THRESHOLD}$. Similarly, B_{LIM} may or may not be used as an operational control point in an MP at which fishing mortality is substantially reduced. Since we are talking about the MSE paradigm in this paper, by reference points we mean the objectives of fisheries management like B_{MSY} that are represented as "known" quantities in Operating Models. To avoid confusion, in the paper we refer to reference points strictly in the MSE interpretation of the term. Quantities that are estimated in an assessment model for the application of a harvest control rule (B_{MSY}, B_0 etc.) are operation control points. Quantities used by tRMFOs exclusively in the Best Assessment context are not considered here.

2. Methods

To organize relevant information on objectives, reference points and PIs, we searched published tRFMO management measures. We considered yield, variability in yield, status and safety objectives, as well as their corresponding reference points and PIs. Note that for ICCAT's northern albacore tuna fishery, there is also a metric for the proportion of years where B_{LIM}<B <B_{THRESH}; because this criterion is not used in the other tRFMOs, we do not explore it here. Similarly, only ICCAT defined a category of performance indicators that are of secondary importance: apart from stability metrics, we filtered out all ICCAT criteria labelled of secondary importance since these cannot be compared among tRFMOs. We included the secondary importance for stability indicators because in this instance we are trying to illustrate the large universe of possibilities that can be used to express this property.

In all cases, broad aspirational objectives like maximizing fishery catch are concretized into PIs. These may be expressed as quantities like mean catch or an analogous quantity like that the stock should be in the green quadrant of the Kobe Matrix with x % probability. Information on the field codes collected in the building of this database is presented in **Appendix 1, Table A1**. Organizing information in this way means that detailed information on management objectives, performance indicators, reference points, probability limits and timeframes for each combination of tRFMO, species, and stocks can be explored in detail.

We present our summary of this information in the broad categories of safety, status, yield, and stability. For each category, we present the diversity of choices made by each tRFMO as frequency histograms of LRPs and yield metrics in **Figures 1-2**. More detailed summaries for safety and stability objectives are presented in **Table 2** and **Table 3**. Because tRFMOs express these probabilities differently, we standardize these probabilities so that they are all expressing the probability of achieving a desired outcome (e.g., biomass is over B_{LIM}) as opposed to the probability of experiencing an undesirable one (e.g., biomass is below B_{LIM}).

3. Results

3.1 Safety

3.1.1 Reference points for safety

tRFMOs have not consistently defined LRPs in the same way (**Figure 1**). These differences are not superficial. The IOTC, IATTC, and the WCPFC have both biomass and fishing mortality limits whereas ICCAT and CCSBT use only biomass LRPs. tRFMOs also express these quantities relative to different metrics. Some use MSY (ICCAT), while some express these limits relative to unfished spawning biomass SSB_0 (WCPFC, IATTC). Still others like the IOTC express LRPs relative to both unfished biomass (B_0) and B_{MSY}.

3.1.2 Performance indicators for safety

While there are different LRPs for safety, the probability requirements for being above these LRPs are relatively consistent. In general, the tRFMOs have stated risk averse probabilities for avoiding LRPs (i.e., that they want to be above them with very high probability) and are more risk neutral for achieving target stock states, consistent with the United Nations Fish Stocks Agreement guidance (UNFSA, Appendix 2, paragraph 5). Adopted probability requirements for being above LRPs ranged from 80 to 90% (**see Table 1),** with the lower probability used for the more conservative LRP of 14% or 20%B_0, which is at or above, respectively, B_{MSY} in those cases.

3.2 Status

3.2.1 Reference points for status

While the tRFMOs shared a commonality in broad terms to be at biomass levels that support MSY based on their Conventions, tRFMOs did not express this broad objective in precisely the same ways. While some tRMFOs express stock status objectives to achieve both B_{MSY} and F_{MSY}, others use only biomass reference points relative to unfished biomass or only fishing mortality reference points.

3.2.2 Performance indicators for status.

While there was variability in how stock status was expressed by tRFMOs, the probability limits for achieving those targets were similar:

- ICCAT: Both the northern albacore and Atlantic bluefin tuna MSEs required a 60% chance of being in the green quadrant of the Kobe plot throughout the projection period.

- IOTC:
  - The bigeye tuna MSE required a 60% probability of achieving the target reference point of SB_{MSY} by 2034-2038.
  - The Swordfish management procedure is designed to achieve a) a 60% probability that the swordfish spawning stock biomass achieves the target reference point of SB_{MSY} by 2034-2038.
  - The skipjack MP was designed with Blim=20% and to have at least a 50% probability that the skipjack tuna spawning stock biomass achieves the biomass level of 40% SB_0 by 2034-2038, which is equivalent to maintaining the biomass above the biomass of SB_{MSY} with 90% probability under the reference set of operating models and 70% under the robustness tests.
- CCSBT: The Southern bluefin tuna MSE required a **50%** probability of achieving a biomass level of 30% of the unfished SSB by 2035. While the required probability is lower than 60%, the target is higher (the SSB_{MSY} proxy is 24%B_0).
- WCPFC and IATTC: The North Pacific albacore MP requires a **50%** chance of having a fishing mortality at or below F45% over the next 10 years. As with CCSBT, while the required probability is lower than 60%, the target is equivalent to an F much lower than F_{MSY} (the SSB_{MSY} proxy is 14%B_0). There is no biomass-based target reference point for the stock. The interim west Pacific Ocean skipjack MP is intended to maintain the stock above the LRP with 80% probability, achieves the objective of relative stability in fishing levels between management periods and in the longer term at roughly 50%B_0.
  .

IATTC agreed to reference points and associated probabilities for the tropical tunas in 2014 in a "best assessment" paradigm, but they are not simulation-tested as MPs, so we do not discuss them here.

3.3 Yield

3.3.1 Performance Indicators for yield

While there was some variability in how tRFMOs define LRPs and status, there was even more variability in the objectives for yield. tRFMOs have variously expressed these objectives in the terms described in **Figure 2**.

3.3.2 Performance indicators for yield

Beyond the value of the catch or mean catches relative to some reference level, we found no probabilistic performance indicators specified for yield metrics. Specifically, they could be expressed probabilistically, for example, as the proportion of years where the catch is above some reference catch level. However, tRFMOs do express some other quantities related to catch performance. For example, a quantity related to minimum catch levels is the probability that the fishery is closed (or the probability that the TAC=0) that is used by ICCAT and the IOTC (**Table 2**). While the WCPFC does not have a minimum catch level, they do have a minimum fishing mortality, connoting the need to maintain some catch levels.

3.4 Stability

3.4.1 Reference points for stability

There is no default reference point like F_{MSY} for variability.

3.2.2 Performance indicators for stability

The most variable performance indicator of all the tRFMOs was variability itself. There were 16 different measures of variability across the tRFMOs (**Table 2**). Some of these quantities are similar in that they express relative changes in total allowable catch (TAC), but others reflect qualitatively different metrics like effort variability and the probability of closures (see above). Currently there are considerably different criteria across the tRFMOs in their MPs for the maximum allowable increase and the maximum allowable decrease in TAC between management cycles (i.e., from one implementation of the MP to the next) (see **Table 3**). For example, the WCPFC had a maximum change in catch from one management cycle to the next of 10% for skipjack, but the WCPFC also designed an MP that had a PI whose results showed a maximum 30% decrease in catch (their management objective 5) from one management cycle to the next.

3.5 Timeframes

The period over which objectives are to be achieved is variable among and within the tRFMOs. How PI statistics are summarized across timeframes is also variable: some PIs are calculated across the entire sample of all OM runs for all years, while other PIs provide statistics across runs that are binned by a single year, and a third possible difference is how PIs are summarized (averaged, etc.) across OMs. But even though the definition of these timeframes is variable in quantitative terms, most have defined short, medium, and long timeframes with a specific set of years over which they are calculated. If applicable, having different timeframes may allow management to visualize any possible tradeoffs across timeframes such as the tradeoff between short term and long-term yield that could, for example, be expected for a rebuilding stock where short term yield might be reduced to allow for higher long-term yields.

4. Discussion and conclusion

While there is a relatively high coherence within a given tRFMO, there is little to no consistency between RFMOs in the specific reference points used or value of PIs. Safety, yield, and variability choices are inconsistent across tRFMOs, and they are sometimes inconsistent within them too. At ICCAT, once the Commission has adopted a set of objectives for one stock/species, they tend to make similar choices for other species, but that is not necessarily the case at other tRFMOs (IOTC, for example).

Since there is no commonly accepted definition of what an LRP is, it is not surprising that each tRFMO selected different values. ICCAT's recommendation 15-07 (1a) defines the LRP as: "…a conservation reference point based on a level of biomass (B_{LIM}) that should be avoided considering that beyond such limits, the sustainability of the stock may be in danger". Management objectives, reference points, and performance indicator statistic thresholds for Safety were less variable within individual tRFMOs but had wider variability across tRFMOs. In some cases, the provenance of LRP choices at tRFMOs is unclear. They appear to reflect the history, culture, and policy of a given tRFMO, as opposed to having a consistent basis in science or international best practice. Apart from the IATTC, we could not find any scientific description for LRP choices at other tRFMOs. There are other possible bases for such choices: various authors have proposed criteria that could justify this choice, like the existence of stationary depensatory stock-recruitment dynamics at low abundance (Liermann and Hilborn 1997, Liermann and Hilborn 2002), non-stationary depensation (Walters and Kitchell 2001), low biomass low productivity state (Forrest *et al*. 2023), or Spawner per Recruit criteria (Mace and Sissenwine 1993). More specific definitions (like avoiding depensation, avoiding recruitment overfishing etc.) about what stock state LRPs are intended to represent might help narrow the variability in tRFMO choices – provided that such quantities could be defensibly estimated. Having LRPs clearly defined would also help limit the choices of probabilities in avoiding them. Formally risk is the product of the probability of an event and the loss given the event; if LRPs represent a point below which a fishery might never recover (depensation) or if they represent recruitment overfishing from which the stock will recover are fundamentally very different decision problems because the corresponding loss functions are very different.

Status reference points were typically related to levels that produce MSY or a proxy thereof. These were expressed in terms of spawning stock biomass in some cases or in terms of B_{MSY} in others; still others used a combination of B_{MSY} and F_{MSY} (the probability of being in the green quadrant of the Kobe matrix, for example). In addition to defined status objectives, yield objectives are also related to B_{MSY} reference points in that many tRFMOs express their catch targets in terms of MSY. For example, the IOTC uses the catch relative to MSY as a performance metric thus implying an objective of achieving MSY, which would require achieving a fishing mortality of F_{MSY} in most cases. Given the convention text of most tRFMOs, and its consistency with the UNFSA, this is hardly surprising.

It is tempting to think that fishing at F_{MSY} will automatically result in a biomass at B_{MSY} and the highest mean catches, but some nuances around B_{MSY} reference points should be considered. A stock at equilibrium with F_{MSY} will also be at B_{MSY}, and assuming suitable selectivity choices, the highest average catches should also be achieved at this same equilibrium. The complication is when the stock is not at equilibrium with the fishing mortality (e.g., due to variable recruitment or variable harvest rates). In such cases, MSY-level catches will not be achieved by fishing at F_{MSY}, and the stock biomass will not be at B_{MSY} either. Accordingly, if the objective is to achieve the highest catches from a stock, then attempting to achieve F_{MSY} may not be desirable. In such cases, non-equilibrium fishing mortality targets, such as X, should be considered.

While there is a lot of variability about how LRPs are defined, yield objectives are relatively consistent. Broadly, these are expressed as some form of maximizing the catch. This was expressed in simple terms like: "maximize average catch;" "maintain catches above average historical catch." But the measures also expressed more nuanced and complex objectives, such as "maintain acceptable CPUE" and "maximize economic yield from the fishery." Some tRFMOs are focused on a single metric (CCSBT, IATTC, ICCAT) but the WCPFC and the IOTC use a suite of catch performance indicators.

Since both objectives and risk tolerance for avoiding an adverse outcome are questions of value, we expected that there would be a diversity of risk tolerance choices that reflect the diversity of people involved. But we did not observe much variability in the diversity of risk tolerances. Instead, this risk tolerance was expressed by using different criteria.

It is important to note that across all species and tRFMOs, the risks that are expressed in PIs are not necessarily equivalent. Firstly, each combination of tRFMO and stock under consideration may compute PIs differently. Consider computing the probability of being in the green quadrant of the Kobe Matrix: some tRFMOs may compute this percentage across all simulations in a given terminal year, or across all years or any number of other possible combinations. While superficially the probabilities might be the same, how these PIs are computed can determine the total risk being expressed. tRFMOs could solve this problem by specifying a common method by which performance statistics are calculated across all stocks. This would ensure coherent interpretability of PI statistics.

While a consistent method for calculating performance statistics would help, there is another factor making apparently identical PI values qualitatively different: the Operating Models. In a given MSE, each Operating Model, or set of Operating Models may make quite different structural (e.g., factors and levels considered in the simulation design) and statistical choices (sampling from the posterior or a multivariate approximation, model weighting). The result of this diversity of choices is that the spread of uncertainty may differ considerably between OMs. This will in turn affect PIs. Consider: OM1 and OM2 are operating models for the same stock; if OM1 and OM2 have the same mean parameters but OM1 has much more variance about the terminal biomass estimate of the conditioning period; correspondingly, it will also have more variability in projected biomass upon the application of a given MP. This effect will be more pronounced at the tails of the distribution of results. When the same MP is evaluated against OM1 vs OM2, the MP evaluated against OM1 will intrinsically have more density below, for example, a LRP like 20% B0 than OM2. Accordingly, that MPs meet the same probability standards (80% chance of being above B_{MSY} for example) does not necessarily mean that the MPs are robust to the same level of risk.

One approach struck us as a practical way to help ensure efficient adoption of management procedures. This is for RFMOs to establish interim reference points and probability requirements for all stocks at the same time. For example, the IOTC's Res. 13-10 and Res. 15-10 establish sets of alternate interim target and LRPs that can behave as defaults until the Scientific Committee advises the Commission of more suitable LRPs. Res. 15-10 further defines alternative interim reference points, defined relative to unfished biomass for those instances where MSY-based reference points cannot reliably be estimated. ICCAT has similar recommendations for Swordfish (Rec. 17-02(6)) and Albacore (Rec. 21-04). The next step after defining default reference points would be to define default probabilities and timeframes over which they are calculated. Such recommendations could be structured in such a way that if more defensible options were determined, then they could replace the default. Default and interim default choices would mean that MP implementation need not be impeded by stock-specific debates about reference point or probability choices.

So, how useful is this survey of tRFMO practices for defining international best practice? Not very. There is too much variability among the tRFMOs to provide extremely precise reference points or probability limit choices. Indeed, rather than illustrate common international best practices, the data collected reveal a large constellation of possibilities. In broad terms, LRPs typically fell out at 40-50% B_{MSY} (ICCAT and IOTC) or 14-20% SSB_0, which is at or above, respectively, B_{MSY} in those cases (WCPFC and IOTC). Further, probability requirements for being above the LRP ranged from 80-90%, with the lower probability used for the more conservative LRP of 20%B_0. For MSE processes in their infancies, parties could explore using the ranges described above but it is not clear that these could be the basis for choices for a given stock or tRFMO. While a common standard for status is more elusive, a minimum 60% probability of being in the Kobe green quadrant is consistent with practice across the tRFMOs.

References

Butterworth, D.S. 2007. Why a management procedure approach? Some positives and negatives. ICES J. Mar. Sci. 64(4): 613–617.

Carruthers, T.R., Huynh, Q.C., Lauretta, M. V., Preece, A.L., Butterworth, D.S., Hordyk, A.R., and Taylor, N.G. 2023. Evaluating Atlantic bluefin tuna harvest strategies that use conventional genetic tagging data. Front. Mar. Sci. 10. doi:10.3389/fmars.2023.1210182.

Cox, S.P., Kronlund, A.R., and Benson, A.J. 2013. The roles of biological reference points and operational control points in management procedures for the sablefish (*Anoplopoma fimbria*) fishery in British Columbia, Canada. Environ. Conserv. 40(4): 318–328. doi:10.1017/S0376892913000271.

Forrest, R. E., Kronlund, A. R., Cleary, J., and Grinnell, M. H. 2023. An evidence-based approach for selecting a LRP for Pacific herring (*Clupea pallasii*) stocks in British Columbia, Canada. Canadian Journal of Fisheries and Aquatic Sciences, 13, 1–13. https://doi.org/10.1139/cjfas-2022-0168

Hall, D.L., Hilborn, R., Stocker, M., and Walters, C.J. 1988. Alternative Harvest Strategies for Pacific Herring (*Clupea harengus pallasi*). Can. J. Fish. Aquat. Sci. 45(5): 888–897. doi:10.1139/f88-107.

Liermann, M., & Hilborn, R. (1997). Depensation in fish stocks: a hierarchic Bayesian meta-analysis. Can. J. Fish. Aquat. Sci., 54, 1976–1984.

Liermann, M., & Hilborn, R. (2001). Depensation: evidence, models and implications. Fish and Fisheries, 2, 33-58.

Ludwig, D., and Walters, C.I.J. 1985. Are Age-structured Models Appropriate for catch and Effort Data? Can. J. Fish. Aquat. Sci. 42(6): 1066–1072.

Mace, P. M., & Sissenwine, M. P. 1993. How Much Spawning per Recruit is Enough? In Risk Evaluation and Biological Reference Points for Fisheries Management. Canadian Special Publication of Fisheries and Aquatic Sciences 120 pp. 101–118. National Research Council of Canada.

Magnusson, A., and Hilborn, R. 2007. What makes fisheries data informative? Fish Fish. 8(4): 337–358. doi:10.1111/j.1467-2979.2007.00258.x.

NRC. 1998. Improving Fish Stock Assessment. National Academies Press, Washington D.C. doi:10.17226/5951.

Walters, C.J., and Parma, A. 1996. Fixed Exploitation Rate Strategies for Coping with the Effects of Climate Change. Can. J. Fish. Aquat. Sci 53: 148–158.

Parma, A.M. 2002. In Search of Robust Harvest Rules for Pacific Halibut in the Face of Uncertain Assessments and Decadal Changes in Productivity. Bull. Mar. Sci. 70(2): 423–453.

Punt, A.E., Dorn, M.W., and Haltuch, M.A. 2008. Evaluation of threshold management strategies for groundfish off the U.S. West Coast. Fish. Res. 94(3): 251–266. doi:10.1109/EMEIT.2011.6023056.

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(2): 303–334. doi:10.1111/faf.12104.

Walters, C.J., and Parma, A. 1996. Fixed Exploitation Rate Strategies for Coping with the Effects of Climate Change. Can. J. Fish. Aquat. Sci 53: 148–158.

Walters, C., & Kitchell, J. F. 2001. Cultivation/depensation effects on juvenile survival and recruitment: implications for the theory of fishing. Can. J. Fish. Aquat. Sci, 58(1), 39–50. https://doi.org/10.1139/f00-160.

Table 1. Summary of the average probability limits for being above LRPs by species and tRFMO. Note that where tRFMOs specify median values for their performance indicators, the reported value is the mean of the medians. CCSBT does not currently have a Limit Reference Point for Southern Bluefin tuna and does not manage any other tuna stocks.

tRFMO	*ALB*	*BET*	*BFT*	*SKJ*	*SWO*	*YFT*
CCSBT	NA	NA		NA	NA	NA
IATTC	80					
ICCAT			85			
IOTC		With high probability		At all times	With high probability	
WCPFC	80	80		80		80

Table 2. List of indicators, including metrics of secondary importance, used to measure stability performance at the tRFMOs.

Performance indicator	*CCSBT*	*IATTC*	*ICCAT*	*IOTC*	*WCPFC*
% Catch coefficient of variation				1	
Average annual variation in catch	1				
Max % change				2	
Mean absolute proportional change in catch			1	1	
Number of TAC changes (count)				1	
Probability of shutdown			1	1	
Probability of TAC change >15%				3	
Probability of TAC change >30%				1	
Probability of TAC change > 10%			1	2	
Probability that a decrease in TAC (or catch for mixed control) is <30% between consecutive assessment periods (once every 3 years), excluding years where TAC=0.		1			
Variability in catch compared to 2013-15 ave. Value of 1 is no variability; value of 0 is relatively high variability in catch					1
Variability in effort. Value of 1 is no variability; value of 0 is relatively high variability in effort					2
Variability in catch			1		1
Variability in effort					1
Odds of no management change		1			
Variability in catch over [x] years			1		

Table 3. Summary of the mean maximum variability in TAC criteria, used within MPs, for increases (VarUp) and decreases (VarDown) between management cycles. The length of the management cycle varies among tRFMOs and species.

Species/Metric	*CCSBT*	*IATTC*	*ICCAT*	*IOTC*	*WCPFC*	*IOTC*
ALB						
Average of VarUp		30%	25%			
Average of VarDown		30%	20%			
BET						
Average of VarUp				15%		
Average of VarDown				15%		
BFT						
Average of VarUp	17.5%		20%			
Average of VarDown	17.5%		35%			
SKJ						
Average of VarUp				26.7%	10%	
Average of VarDown				27.5%	10%	
SWO						
Average of VarUp						10%
Average of VarDown						15%

*Southern bluefin tuna has a maximum TAC change of 3000 t, which is approximately 17.5% of the current TAC.

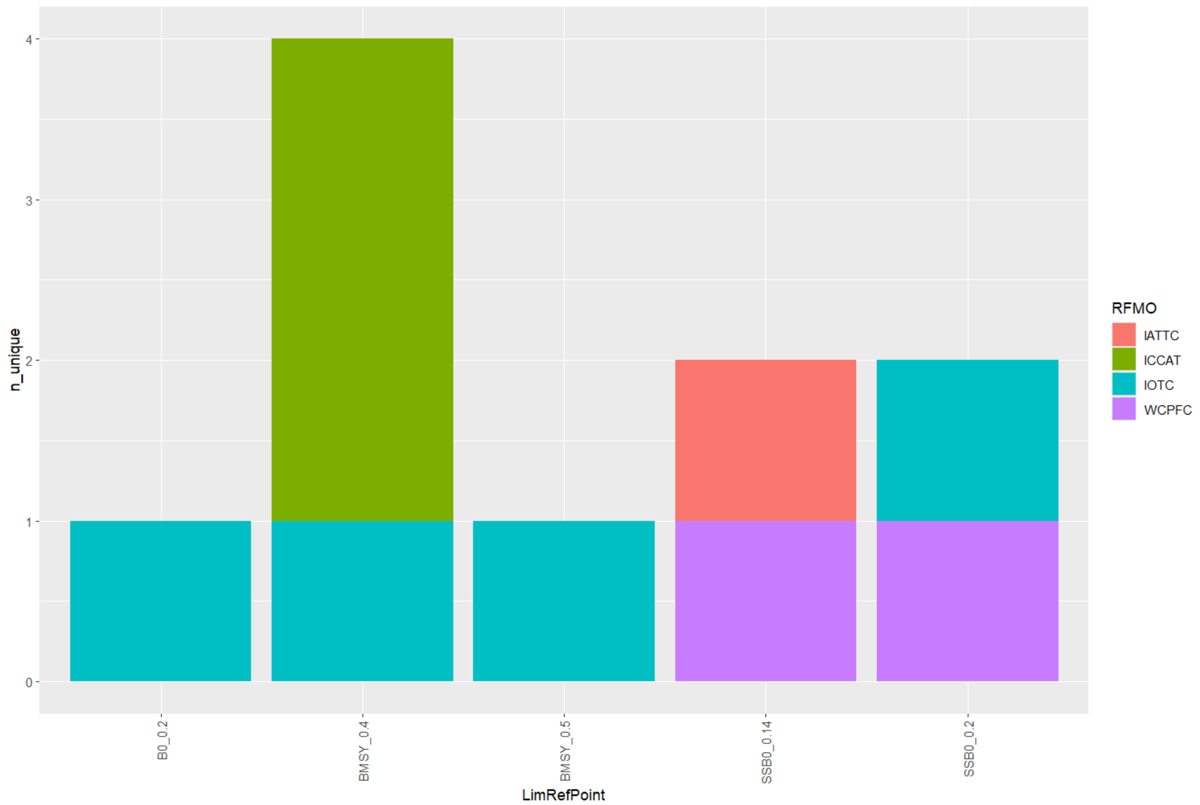

Figure 1. Sum of the unique counts (n_unique y) of LRPs (x) used by tRFMOs. B0_0.2 represents 20% of the unfished biomass, BMSY_0.4 is 40% of B_{MSY}, BMSY_0.5 is 50% of B_{MSY}, SB0_0.2 is 20% of the unfished spawning stock biomass, and SSB0_0.14 is 14% of the unfished spawning stock biomass. North Pacific albacore spans both IATTC and WCPFC and both tRFMOs adopted SSB0_0.14 as the LRP for North Pacific albacore; so, while there are two occurrences in this figure of SSB0_0.14 they both relate to the same stock.

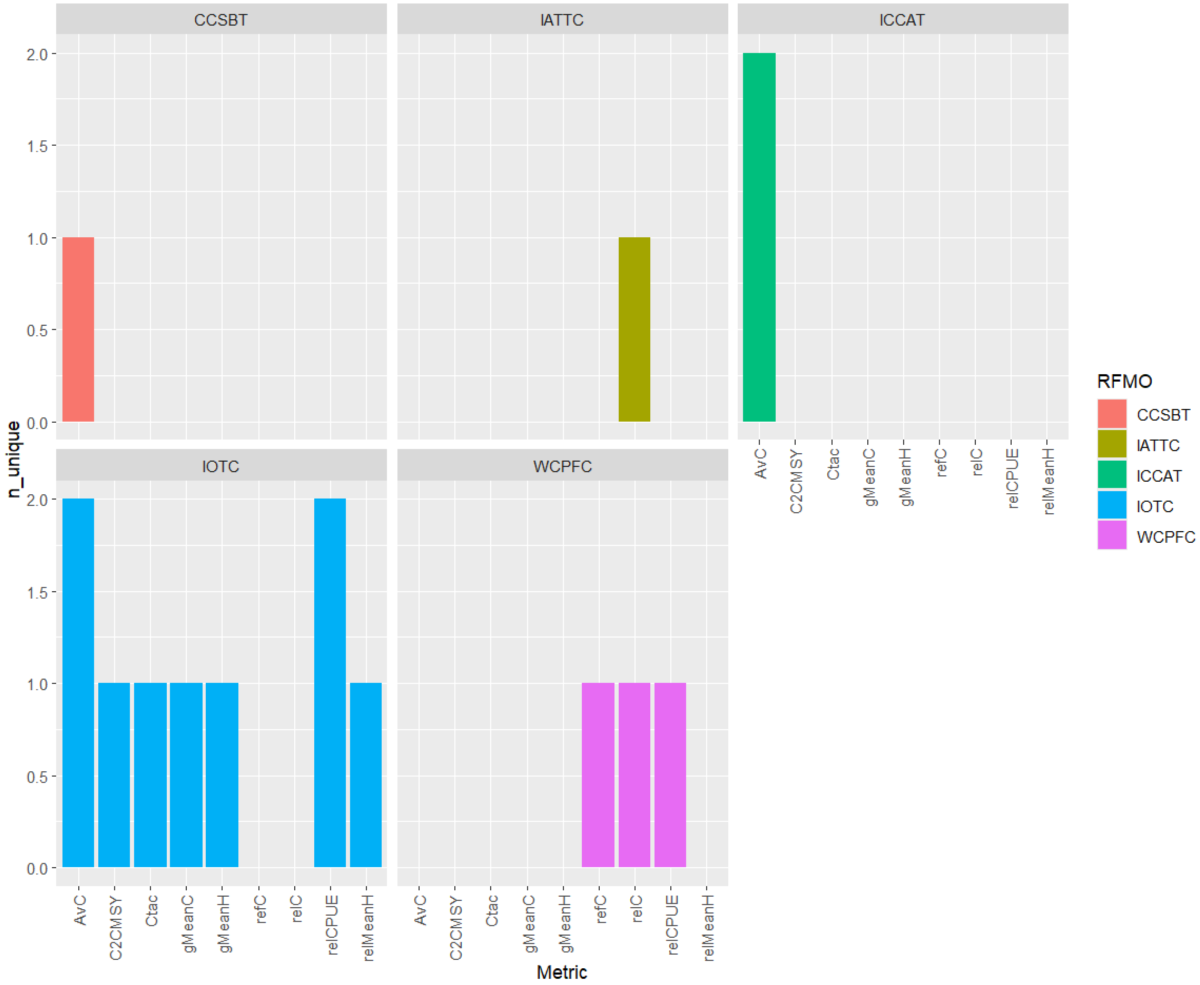

Figure 2. Sum of the unique count (y) of yield metrics (x) for each tRFMO (across all time periods). AvC represents the average catch, C2CMSY is the ratio of the catch to the catch that could be achieved at B_{MSY}, Ctac is the mean total allowable catch, gMeanC is geometric mean of the catch, gMeanH is the geometric mean of the harvest rate, refC is the mean of catch from 2013-2015, relC is the probability that catch in any given year of the MSE forward simulation is above average historical (1981-2010) catch, relCPUE is CPUE for pole and line fisheries relative to 2001-04 average, relMeanH is the relative exploitation rate (geometric mean).

Appendix 1

Table A1. Fields and field definitions used for defining tRFMO reference point and performance indicators.

Field Code	Field Description
RFMO	The regional management organization responsible for the stock
Species Group	The broad taxonomic category
Species	The common name of the species
Stock	The management unit within the RFMO
Min Year	The first year of the time interval over which a performance indicator is calculated
Max Year	The last year of the time interval over which a performance indicator is calculated
Timeframe	The broad category of the time frame (short term, medium term, long term)
Metric Category	Status, safety, stability, or yield
Corresponding management objective	Text from a measure defining this management objective
Metric	Name or code for the PI
Metric Value (where applicable)	Probability of reaching a target, limit, or max variance value
Metric Description	Description in words (quote from recommendation/resolution if possible)
Metric Priority	Primary, secondary or tertiary
Related Reference Point(s)	Code of symbol for related quantities
RelativeReference	If applicable, where the reference point is expressed as a fraction of another e.g., 0.2 for B_{MSY}
RefPointFrac	The fraction of the relative reference
AdoptedStatus	If the PI or reference point has been officially adopted by the RFMO or not
Probability Limit Adoption status	If the probability limit has been adopted or not
Reference(s)	Reference/hyperlink tRFMO measure
Notes	