THE POTENTIAL OF CONVENTIONAL GENETIC MARK-RECAPTURE FOR INFORMING MANAGEMENT PROCEDURES AND STOCK ASSESSMENTS FOR ATLANTIC BLUEFIN TUNA

T. R. Carruthers[1] and Q. Huynh[1].

SUMMARY

An individual gene tagging model was developed and embedded within the operating models of the existing Atlantic bluefin tuna MSE framework. A multi-year tagging estimator was developed, and its estimation performance evaluated over multiple gene tag release distributions, release numbers and fishery exploitation rates. The tagging estimator was used in a constant harvest rate management procedure and its performance evaluated using operating models from the bluefin tuna MSE framework. Tag release designs that followed the historical pattern of electronic tagging were the most appropriate of those evaluated, offering precise estimates of exploitation rates for both western and eastern stocks. Effective releases of 400-500 per year were sufficient to obtain relatively precise estimates of exploitation rate after 4 release years. Management procedures using the tagging estimator had expected performance that was comparable to idealized 'perfect information' constant exploitation rate management procedures. In general, conventional gene tagging is a promising basis for calculating management advice for Atlantic bluefin tuna that is cheaper, simpler and more robust than the conventional stock assessment paradigm.

RÉSUMÉ

Le présent document fournit une mise à jour de l'étude sur l'utilisation de l'habitat de l'espadon, développée dans le cadre du plan de travail du Groupe d'espèces sur l'espadon de l'ICCAT. Un total de 26 marques miniPAT ont été déployées jusqu'à présent dans l'Atlantique Nord (n=13) et Sud (n=9) et en Méditerranée (n=4). Les données de huit marques ont été analysées afin de déterminer l'utilisation horizontale et verticale de l'habitat. Ces résultats préliminaires ont montré que l'espadon se déplaçait dans plusieurs directions, parcourant des distances considérables dans les stocks nord et sud. L'espadon a passé la majeure partie de la journée dans des eaux plus profondes avec une moyenne de 540,8 m, étant plus proche de la surface pendant la nuit (moyenne=78,3 m). La plongée la plus profonde enregistrée était de 1.480 m. En ce qui concerne la température, l'espadon habite des eaux dont la température varie de 3,9ºC à 30,5ºC avec une moyenne de 11,3ºC pendant le jour et de 21,7ºC pendant la nuit. L'objectif principal de la prochaine phase du projet est de poursuivre le déploiement des marques en 2022 dans plusieurs régions de l'océan Atlantique et de la mer Méditerranée. Actuellement, 11 marques ont été distribuées aux CPC participantes et neuf marques doivent encore être attribuées.

RESUMEN

Este documento proporciona una actualización del estudio sobre el uso del hábitat del pez espada, desarrollado dentro del plan de trabajo del Grupo de especies de pez espada de ICCAT. Hasta ahora se han colocado un total de 26 marcas miniPAT en el Atlántico norte (n=13) y sur (n=9) y en el Mediterráneo (n=4). Se analizaron los datos de ocho marcas para determinar el uso horizontal y vertical del hábitat. Estos resultados preliminares mostraron que el pez espada se movía en varias direcciones, viajando distancias considerables tanto en los stocks del norte como en los del sur. El pez espada pasó la mayor parte del día en aguas más profundas, con una media de 540,8 m, y permanecía más cerca de la superficie durante la noche (media de 78,3 m. La inmersión más profunda registrada fue de 1.480 m. En cuanto a la temperatura, el pez espada habitaba en aguas con temperaturas que oscilaban entre los 3,9º C y los 30,5º C, con una media de 11,3º C durante el día y 21,7º C durante la noche. El plan principal para la próxima fase del proyecto es continuar el despliegue de marcas durante el año 2022 en varias regiones del océano Atlántico y el mar Mediterráneo. En la actualidad, las CPC participantes cuentan con 11 marcas y quedan nueve por atribuir.

[1] Blue Matter Science Ltd., 2150 Bridgman Ave, North Vancouver, B.C. V7P 2T9. tom@bluematterscience.com

KEYWORDS

Management Strategy Evaluation, gene tagging, mark-recapture, stock assessment, bluefin tuna, management procedure

Introduction

Atlantic bluefin tuna presents a formidable fishery management challenge. Coordinated by the International Commission for the Conservation of Atlantic Tunas (ICCAT), the primary management measures are annual total allowable catches (TACs) for East and West management areas. However, the Atlantic population is known to consist of at least two genetically distinct but geographically overlapping stocks (Block et al. 2015, Puncher et al. 2018, Rooker et al. 2014) that are at least an order of magnitude different in size (ICCAT 2017, 2021) and mix in the East and West management areas (**Figure** 1). The much larger eastern stock spawns in both the western and eastern Mediterranean (Alemany et al. 2010, Hernandez et al. 2022); the smaller western stock spawns in the Gulf of Mexico and the Slope Sea (Richardson et al. 2016). The distribution of the stocks, and hence their mixing, is seasonal and the degree of movement varies depending on the size of individual fish (Fukuda et al. 2010; Galuardi and Lutcavage 2012, Lutcavage et al. 1999). Superimposed on these complex stock dynamics are a range of fisheries, diverse in their geographic range and the size of fish that they catch (ICCAT 2017, 2021). For example, longline fleets are highly mobile and can selectively fish throughout the North Atlantic while Mediterranean trap fisheries typically operate from a static location.

ICCAT specifies TACs for the East and West management areas using discrete area-specific stock assessments that aim to characterize population and fishery dynamics in order to quantify population abundance and sustainable exploitation rates (ICCAT 2017, 2021). There are several complications in the interpretation of stock assessment advice: (1) the assessments are defined geographically by management area and do not necessarily reflect stock dynamics that include mixing, particularly of eastern fish into the West management area; (2) it has been increasingly difficult to establish scientifically defensible stock assessments (Maunder 2021); (3) despite extensive study there remains considerable uncertainty over fundamental biological, ecological and life-history dynamics such as natural survival rate and the maturity-at-length schedule (Diaz and Turner 2007).

To address these challenges, ICCAT has pursued Management Strategy Evaluation (MSE) (Butterworth and Punt 1999, Cochrane et al. 1998) as an alternative framework which aims to identify simple management procedures (algorithms for setting TACs, aka management procedures) that are robust to the full range of uncertainties, many of which currently prevent the identification of a single defensible stock assessment (GBYP 2019). Such empirical management procedures typically make use of one or more indices of relative abundance which are expected to track vulnerable biomass or spawning stock biomass (e.g., Carruthers et al. 2015).

A principal concern in establishing such management procedures for bluefin tuna is the availability of suitable data. Given the existing conflicts among regional indices of relative abundance (including those based on commercial catch rates and fishery-independent surveys), and the concerns of various scientists and stakeholders over the representativeness of particular abundance indices, a lack of suitable data remains a potential obstacle to long-term adoption of a management procedure for Atlantic bluefin tuna.

In parallel to MSE development, a close-kin gene tagging program has been initiated in the West Atlantic (also known as close-kin mark recapture or CKMR). In essence, CKMR is intergenerational mark recapture that aims to estimate the size of the spawning stock biomass based on the ‘recapture’ of genetic markers sampled on juveniles. Although it is a new development and has yet to provide estimates of western spawning stock biomass for Atlantic bluefin tuna (Grewe et al. 2018), a successful program has been in place for Southern bluefin tuna for more than 15 years (Bravington et al., 2016). The viability of the approach is being evaluated for the eastern Atlantic stock (Anon 2021a).

While CKMR is a promising avenue for informing management procedures, there are a number of potential concerns with applying the approach to Atlantic bluefin tuna. The first is that for a given stock of bluefin tuna it can be costly and requires many thousands of juveniles and adults to be genotyped to obtain the number of parent-offspring matches required to precisely quantify spawning biomass. Depending on abundance and dynamics it may not be possible to obtain precise estimates of spawning stock biomass. Another concern is that close-kin only provides estimates of spawning biomass which in the Atlantic does not necessarily correspond with management area and therefore may not provide a reasonable basis for informing area-based management decisions. The third

concern is that close-kin has not been subjected to rigorous theoretical testing. For example, given heritability in recapture probability (e.g., parents and offspring are geographically correlated, sub-stock structure), CKMR may provide biased estimates of spawning stock biomass (which fortunately can be detected from the data if it occurs). Lastly, close-kin programs are typically regional and exclude the full range of contracting parties in the collection of data. It follows that they may be perceived as a program promoted by a subset of stakeholders.

The focus on CKMR leap-frogged an older and more established approach for estimating abundance and exploitation rates: conventional gene tagging (Andreou et al. 2011; Lukacs and Burnham, 2005; Preece et al. 2015, 2018) in which fish are genetically sampled at sea, released and then recaptured. From such data, well-established mark-recapture estimators of abundance and exploitation rate can be applied (e.g., Cormack 1964).

For many reasons, conventional gene-tagging is an outstanding opportunity not only for bluefin tuna management in the Atlantic but also globally. As with CKMR, conventional gene tagging obtains essentially 100% reporting rates and 0% shedding rates. Atlantic Bluefin tuna tagged in-water or by rod and reel, exhibit close to zero post-release mortality (Marcek and Graves 2014, Stokesbury et al. 2011) which can otherwise add an additional source of uncertainty and increase both bias and imprecision of abundance estimates. Roughly 85% of global bluefin tuna are sold in a handful of Japanese markets, offering a centralized and convenient location for genetic 'recapture'. The cost of a gene tag recapture relative to the value of a bluefin tuna is very low. Preece et al. (2018) estimated the cost of genotyping to be approximately US$15 per sample including DNA extraction, equipment and labor. In the same year as that publication, Atlantic bluefin tuna had an average dock price of US$12,000 per tonne and a mean end value of US$37,000 per tonne (Pew 2020). Consequently, while releases are expensive (at-sea sampling in addition to the opportunity cost of returning a fish), the cost of tagging a fish caught for consumption is essentially negligible (less than half a percent of the end value of a fish). This opens the potential for genetic recapture of a large fraction of market fish which in turn would require a relatively small number of in-water releases to obtain sufficient precision in estimates of abundance and exploitation rate. Such a program could be extensible to all ocean areas, including the Pacific and Southern Oceans (e.g., Bradford et al. 2015). The exploitation rate estimates arising from such a recapture program can feed directly into a management procedure for West and East Atlantic bluefin tuna (also Pacific, and Southern stocks) bypassing stock assessment processes that have been increasingly problematic over recent years. A similar gene-tagging program for monitoring abundance of juvenile southern bluefin tuna was implemented in 2016 and hence many of the technical challenges of genotyping bluefin have been overcome. Additionally, if evaluation of stock status is necessary, conventional gene tagging can provide estimates of exploitation rate and abundance that can directly inform stock assessments. Once established over multiple years, conventional gene tagging can also provide estimates of natural survival rate, which is a key source of uncertainty in current stock assessments for Atlantic bluefin tuna.

Management procedures that provide advice based on trends in exploitation rate and abundance can provide excellent management performance that is not necessarily affected by persistent biases in estimation of these quantities. Since they do not rely on many assumptions about stock and fishery dynamics, management procedures using direct information from conventional gene tagging can be expected to be robust to a wide range of uncertainties regarding stock mixing and migration. Conventional gene tagging data have already been used in a management procedure for Southern bluefin tuna (Preece et al 2020, Hillary 2019). In theory, the cost of such a program for Atlantic bluefin tuna would be of the same order of magnitude as current data collection and assessment frameworks and could provide more accurate and precise estimates of quantities informative for decision making. Additionally, gene tagging is a collaborative endeavor and all contracting parties have a role in the collection of data used to inform management.

Before such a substantial shift in management paradigm can be recommended, first there must be a suitably rigorous test of the approach; that the theory proposed can be expected to provide suitable performance given the various uncertainties in bluefin tuna population and fishery dynamics. The MSE framework for Atlantic bluefin tuna is ideally suited and includes a wide range of operating models that represent plausible alternative states of nature for the mixed stock population and fishery dynamics (Anon 2022, Carruthers et al. 2015, 2020, Carruthers and Butterworth 2017, ICCAT 2020).

Using the peer-reviewed MSE framework for Atlantic bluefin tuna we aim to evaluate the expected precision and bias of genetic mark-recapture estimates of exploitation rate (and abundance) given varying distributions of gene tag releases, release numbers and fishery exploitation rates. Furthermore, we quantify the expected management performance of management procedures based on mark-recapture estimators informed by conventional gene tagging.

Methods

Operating Models

The existing MSE framework for Atlantic bluefin tuna includes 48 seasonal, spatial, multi-stock operating models designed to span uncertainty in biological and fishing dynamics (Carruthers 2020). In this simulation evaluation we focused on the 'Reference Case' operating model used as the primary basis for preliminary investigations in the MSE process, in which both western and eastern stocks are assumed to have relatively low stock sizes (Low abundance). Two alternative operating models are also considered that span more extreme scenarios for stock mixing where the eastern abundance was relatively high (High eastern) and both eastern and western stock abundance is high (High abundance) (**Figure 2**). A total of 150 simulations were conducted for each operating model including stochasticity in recruitment error and the distribution of tag releases. This level of replication provided stable (within a percent) estimates of precision and bias in exploitation rate by tagging estimators and also stable estimates of management performance of management procedures using those estimators.

To generate simulated tagging datasets, a range of simulated exploitation rates (the ratio of catch biomass to vulnerable biomass) were projected (2%, 4%, 6%, 8% and 10%) to evaluate the precision of the tagging estimator given varying recapture probabilities due to fishery exploitation level.

Individual-Based Tagging Model

An individual-based tagging model (IBTM) was coded into the Atlantic Bluefin Tuna MSE framework (for a complete mathematical description of the model see Appendix A). In this model, the dynamics of the tagged fish follow those of the 7-area, seasonal multi-stock operating model that simulates age vulnerability to fishing, movement and mixing of the untagged population. The IBTM is stochastic with respect to tag release distribution, age-specific movement, natural survival and harvesting.

Non-independence in recapture probability among tags is captured by shared seasonal movement and seasonal-spatial patterns of fishery exploitation. No additional sources of non-independent in recapture probability were simulated (e.g., schooling of tagged fish and subsequent heterogeneity in movement dynamics). The model assumes that tagged and untagged fish have equal recapture probability. Older technologies such as fin-clipping could leave visible evidence of a tag release which could affect recapture probability, however residual trauma is almost undetectable for newer gene tagging technologies (e.g., Bradford et al. 2015). The model assumes that stock of origin (eastern / western stock) is accurately assigned to fish. This rate of error is likely to be low and, in any case, ignorable in situations where there are approximately comparable recapture numbers of tags released on eastern and western fish.

The IBTM records the fleet that recaptures each tag allowing for varying genotyping rate at recapture among fleets (e.g., varying fractions of recaptured fish at a Japanese market). In the absence of reliable data on the fate of fish caught by each fleet, this analysis remains generic and assumes homogeneity in recapture probability among fleets.

Genetic cross-contamination and gene tagging errors were assumed to be ignorable. In reality these errors would reduce the recapture probability and could be captured in the 'natural survival' rate parameter of a multiyear tagging estimator and in any case would not affect trends in estimated exploitation rate. If such errors could be independently quantified these could be added to a tagging estimator to adjust for the downward bias in estimated exploitation rate.

Tag Release Distributions

The simulation model includes nine fleets, seven areas and four seasons (a total of 252 strata). Six tag release distributions were considered in this simulation evaluation, all of which are considered practically feasible as they are based on the spatiotemporal distribution of either observed catches, previous tagging release programs or fishing effort:

1) Releases in proportion to catches. In this design annual tag releases are assumed to occur in proportion to the catch distribution by weight in the most recent historical year of the simulation (2019).
2) As #1 but balanced such that the same number of tags are released in the East and West areas.
3) Uniformly distributed across strata with positive catches (approximately 10% of fleet-area-season strata recorded catches in 2019)
4) In proportion to estimated exploitation rate in 2019 (as estimated by the Reference Case operating model).

This is intended to broadly represent the distribution of recent fishing effort.

5) In proportion to the distribution of historical conventional tag releases.
6) In proportion to the distribution of historical electronic tag releases (Anon 2021b).

These release distributions vary markedly in their distribution among fleets, areas and seasons (**Figure 3**). Theoretical release distributions that could not be implemented in practice were not evaluated such as tags released in proportion to simulated abundance.

Release Numbers and Recapture Probability

Three levels of tag release frequency were evaluated: 100, 200 and 500 tags per year. In reality it is not possible to genotype every fish that is caught - the genotyping rate at capture would be less than 100%. Since the probability of recapture is proportional to the number of tags released and the genotyping rate at capture, it is trivial to infer the equivalent release numbers required to obtain comparable estimation performance with a lower genotyping rate at capture. For example, the 100, 200 and 500 tag releases per year given 100% genotyping rate are equivalent to 500, 1000 and 2500 releases per year given a 20% genotyping rate at capture. For computational efficiency we evaluated precision and bias in tagging estimators assuming 100% genotyping rate at capture and later discuss the implications for release numbers given lesser and more plausible genotyping rates of recaptured fish. Note that this assumes that the genotyping rate at capture can be accurately quantified, which is likely given the comprehensive monitoring of overall bluefin tuna catch numbers. Herein, we refer to the number of annual releases equivalent to a genotyping rate at capture of 100% as 'effective releases' or 'effective release numbers'.

Mark-Recapture Estimator

A multi-year Brownie estimator was developed in order to calculate exploitation rate estimates sufficiently efficiently to be used within a management procedure (see Appendix B for details). The estimator calculates annual exploitation rates and a parameter that adjusts for higher expected mark rates in the initial release year due to incomplete mixing / seasonal releases. The model estimates both the maximum likelihood estimate of the annual exploitation rates in addition to their precision and covariance. In this analysis the focus is on management procedure performance which, as explained below, is invariant to biases in the mark recapture estimator. For this reason natural survival was assumed to be known and not estimated.

Management Procedures

Management procedures assumed a two-year management interval (TACs are updated every two years) and MSE closed-loop projections were carried out over a 20 year projection (2023-2042). A release program was assumed to have occurred in the three years prior to the first projection year (2020-2022). Management procedures that use the multi-year Brownie estimator varied in their target exploitation rate and were tested for annual effective release numbers of 200 and 500 tags. The gene tagging management procedures adjusted the TAC in a management area according to the ratio of recent (mean over n years) stock-specific exploitation rate estimates (U) to a target exploitation rate (T):

$$TAC_y = TAC_{y-1} \cdot nT / \sum_{y-1-n}^{y-1} U_y \quad \text{(Eqn. 1)}$$

In these investigations the mean estimated exploitation rate of the last two years (n = 2) was used for TAC calculation. Management procedures were specified in pairs with the same target exploitation rate for both stocks. Performance of the tagging management procedures was framed by a range of constant exploitation rate management procedures (also paired with the same exploitation rate in each management area) based on perfect information of vulnerable biomass in the management areas. In practice, management procedures are rarely implemented without limits on the maximum change in TAC among management updates. All management procedures were subject to a relatively permissive maximum upward and downward adjustment of 50% between management updates.

Results

Estimation Performance Given Alternative Tag Release Distributions

The tag release distribution strongly determined both the accuracy and precision in the Brownie model estimates of exploitation rate (**Figure 4**). Since overall catches are biased strongly towards the eastern stock (**Figure** 3, top left panel) relatively few tags are released on western fish with this release distribution and hence there is a high degree of error in maximum likelihood estimates of exploitation rate (**Figure 4**, panel a), and those estimates have relatively high variance (i.e., coefficient of variation greater than 0.5, **Figure 4**, panel b). Altering this release distribution to ensure an equal number of releases in the West and East management areas dramatically reduces error and variance in estimates for the western stock while largely unaffecting estimation performance for the eastern stock (**Figure 4**, panels e-h).

Both of these 'proportional to catch' release designs led to positively biased estimates of exploitation rates, generally overestimated by between 15-50% for both western and eastern stocks suggesting unrepresentative mark rate due to incomplete mixing of tags throughout the vulnerable population. This heterogeneity in mark rate with higher rates in areas of exploitation is expected since release and recapture of tags occurs on fish in specific locations and sizes classes. When tags were released uniformly across areas, fleets and seasons where positive catches are observed, estimates of exploitation rate were much less biased and unbiased in some recapture years (**Figure 4**, panels i-l).

Releasing tags in proportion to exploitation rate (as a proxy for fishing effort) provided comparable estimation performance to when tags were released in proportion to catch but balanced in the East and West, but estimates were somewhat more positively biased and imprecise (**Figure 4**, panels m-p).

Tag releases that were distributed according to historical conventional tag releases led to insufficient recaptures to estimate exploitation rate for the western stock using the Brownie model. The estimates of eastern exploitation rate were however the least biased of all release distributions.

Tag releases that were distributed according to historical electronic tag releases produced the most precise estimates of western stock exploitation rate by some margin with CV in estimates of around 0.1 to 0.13 in years 2016-2018 (**Figure 4**, panel t) compared with 0.2 to 0.19 for the uniform release distribution over strata with positive catches (**Figure 4**, panel j). There was not an appreciable impact on exploitation rate estimation for the eastern stock. This distribution of tag releases produced exploitation rates for both stocks that were positively biased between 10% and 30% for years 2026-2028, demonstrating that mark rates experienced by fishing are a commensurate fraction higher than the vulnerable stock in general. This release distribution was used in all analyses herein because precision in estimates (reported CV of the estimator) is a greater determinant of management procedure performance than a consistent bias in estimates. Management procedures typically include tuning parameters that control the aggressiveness of the management procedure (i.e. catch relative to biomass performance). These tuning parameters are adjusted to obtain specified biological or yield performance allowing candidate management procedures to be compared while controlling for one axis of the prevailing yield-biomass trade-off. In a management procedure that aims for a fixed exploitation rate, the tuning parameter is typically the target exploitation rate. It follows that precision in exploitation rate estimates controls the responsiveness of the strategy to changes in simulated exploitation rate and is therefore critical in MP performance. However consistent biases are ignorable because these are simply countered by a commensurate increase in the target exploitation rate tuning parameter.

Estimation Performance for Various Exploitation Rates and Effective Release Numbers

In general, the effective number of tag releases did not impact the accuracy of exploitation rate estimates using the Brownie model (i.e. **Figure 5** where the solid colored points in the 'MLE' panels were comparable in level among the colors that represent release numbers). Given releases proportional to the distribution of electronic tagging, exploitation rate estimates were generally positively biased to approximately the same degree independent of the simulated fishery exploitation rate (approximately 15-40% in years 2024 - 2028, **Figure 5**). In most combinations of exploitation rate and release numbers, precision in estimates of exploitation rate declined to an asymptote by recapture year 6 (2028) (e.g., **Figure 5**, panel j) with release number having a stronger impact on precision than simulated exploitation rate.

The precision of the Brownie estimator, represented as a coefficient of variation σ, was well approximated by a linear model in which the inverse square-root of exploitation rate u, and release numbers n, were explanatory factors (i.e., following the expected relationship between standard error and sample size) (see **Figure** 6 for estimated precisions versus the linear model prediction of precisions):

$$\sigma_i = \alpha \frac{1}{\sqrt{n_i}} + \beta \frac{1}{\sqrt{u_i}} + \gamma \frac{1}{\sqrt{n_i}\sqrt{u_i}} + \varepsilon_i \quad \text{(Eqn. 2)}$$

Deriving this relationship allows for the production of tables of the required number of annual effective releases (annual release numbers multiplied by the genotyping rate) given a specified estimation precision (Table 1). Due to mixing of eastern fish in the West area, a much larger number of tag releases are required to obtain comparable precision of exploitation rate estimates of western fish (Table 1).

Implications for 'Optimal' Tagging Design

The precision of the tagging estimator is non-linear with respect to the effective number of annual releases which is the product of annual release numbers and genotyping rate. The cost of release and genotyping programs may be expected to be approximately linear. It follows that if the costs of genotyping C_{gen}, and tag releases C_{rel}, are available, for a given exploitation rate u, desired level of precision in the tagging estimator σ, and the number of annual catches n_{caught}, it is possible to numerically solve for the least expensive combination of tag release numbers nrel, and genotyping rate G (using the linear model and fitted parameters α, β, γ of Eqn 2):

$$\min_{n_{rel},\ G} \quad \left(n_{caught} G C_{gen} + n_{rel} C_{rel}\right) \left[\sigma - \left(\alpha \frac{1}{\sqrt{n_{caught} G}} + \beta \frac{1}{\sqrt{u}} + \gamma \frac{1}{\sqrt{n_{caught} G}\sqrt{u}}\right)\right]^2$$
$$s.t. \quad 0 < G \leq 1$$
$$0 < n_{rel} \quad \text{(Eqn. 3)}$$

To demonstrate the sensitivity of 'optimal' tagging designs (release numbers and genotyping rate) to costs and desired precision in the estimator, a set of default values were assumed for the desired precision (σ = 0.2), fishery exploitation rate (u = 4%), number of annual fish caught and landed (n_{caught} = 250,000, following estimated catch numbers from the operating model in 2019), cost of releasing a tag (C_{rel} = \$1,000 per tag, given at-dock price of \$12,000 per tonne, Pew 2020), cost of genotyping a fish at recapture (C_{gen} = \$15, Preece et al., 2018), and each of these parameters were then systematically varied (Table 2). Regardless of the values of these parameters, 'optimal' tagging designs always invest substantially more in the recapture program (genotyping rate) than the release program, typically by a factor of between 2 and 7 depending on parameter values (Table 2). 'Optimal' genotyping rates were generally higher than 70% - the large majority of captured fish. At default settings for model parameters the cost of the program was estimated at approximately US\$3.5m if focused on the precision of the western exploitation rate and \$3m if focused on precision in eastern exploitation rate estimates. The optimal genotyping rate and release number (hence effective release number) were essentially invariant to the number of fish caught, the release cost and the genotyping cost (Table 2). It follows that only the target exploitation rate and the desired precision of the estimator strongly impact the optimal tagging design.

Effect of Operating Model on Estimation Performance

In general, the degree of stock mixing and the relative magnitude of the two stocks did not substantially impact the precision of the tagging estimator (CV) but could impact the degree of positive bias (**Figure** 7). The 'High abundance' operating model where the magnitude of both stocks was larger was around 15% more positively biased for both stocks than the 'Low abundance' operating model for the western stock in 2028 (**Figure** 7). The estimation performance of the tagging model was more comparable among the 'High eastern' and the 'Low abundance' operating models.

Management Procedure Performance

The simulated conditions of the reference case 'Low abundance' operating model simulates relatively high recruitment and large vulnerable biomass, particularly for the eastern stock. Hence TAC adjustments of the 'Perfect Information' management procedure were generally increasing for the first 5-10 years after which they declined and stabilized (**Figure** 8).

Projected yields for gene tagging management procedures tended to lag the perfect information management procedure by around 5 years with catches increasing more slowly in response to larger available vulnerable biomass, and then also declining later as vulnerable biomass declines (**Figure** 8). At higher target exploitation rates (i.e. above 6%), there was a much wider discrepancy in projection among gene tagging and perfect information management procedures. Catches of the gene tagging management procedures tended to increase to much higher levels leading to biomass trajectories that were in steep decline at the end of the 20-year projection (**Figure** 8 panels c and d).

For all management procedures, the projected pattern of catches and biomass were similar for the East area / eastern stock and the West area / western stock (**Figure** 8). In general, the pattern in catches and biomass outcomes was very similar among the gene tagging management procedures independent of the number of effective annual tag releases (**Figure** 8). However there was greater variability in both annual yield and biomass outcomes given the lower level of 200 effective tag releases (**Figure** 9).

In fisheries MSE in general, the most prevalent management performance trade-off among management procedures of varying type and parameterization exists between the yields (catches) that are taken and the biomass that remains after these are removed. Overall yield performance of the management procedures was defined here as the mean annual catches in the first 20 projection years (2023 - 2042). Biomass performance was defined as stock biomass relative to BMSY in the final projection year (2042). The expected (mean) yield and biomass performance of the gene tagging management procedures was similar irrespective of the number of annual effect tag releases (**Figure** 10). Surprisingly there was a relatively small expected gain in yield from management using the idealized fixed exploitation rate strategy ('Perfect Information'). In the West and East management areas the expected gain in yield when moving to 'Perfect Information' was less than 10% and 20%, respectively.

Much larger differences could be seen in the lower tail of yield and biomass outcomes (i.e. the 5th percentile of outcomes, **Figure** 10). In both management areas and stocks, gene tagging management procedures had much more uncertain yield and biomass outcomes than the reference 'Perfect Information' management procedure. Tag release numbers could strongly impact lower tail outcomes for yield in the West area and biomass of the eastern stock. In the West area, releasing 200 effective tags per year resulted in a 5th percentile of yields that was half that of simulations with 500 effective tags per year (**Figure** 10, left hand panel). In contrast, eastern stock biomass outcomes could drop close to zero (stock extirpation) for the higher target exploitation rates in simulations with only 200 effective tag releases per year (**Figure** 10, right hand panel).

At the higher target exploitation rate levels there were substantially worse biological outcomes for the western stock (with more steeply declining biomass trends, **Figure** 8). Gene tagging management procedures with target exploitation rates of 8% and 10% differed strongly from the Perfect Information management procedure, catching 50% more and depleting the western stock to below half of BMSY after 20 projection years (**Figure** 8).

Discussion

Developing an individual tagging model embedded within the operating models of the Atlantic bluefin tuna MSE framework allowed the performance of a gene tagging estimator to be evaluated subject to complex stochastic seasonal and spatial multistock population dynamics. These include stock-, season- and age-specific movement and varying stock mixing depending on the relative abundance and age structure of the East and West stocks. In these simulations, tags were released and recaptured from fishing fleets for which spatio-temporal distribution and size-selectivity are estimated from empirical data. It follows that the approach rigorously accounts for dynamical processes that can create non-independence in recapture probability. For example, elevated higher mark rates in high exploitation rate areas, and comparable movement of tags released on similar size fish in the same season and area. The stochastic nature of the simulations captures other important phenomena that could affect tagging estimators such as increased tag mixing (homogeneity in mark rates) with tag time-at-liberty.

Using these simulations it was possible to quantify the resulting bias and precision of a multiyear exploitation rate estimator given varying distributions of tag releases, fishery exploitation rate and effective release numbers. At intermediate fishery exploitation rates of 6%, effective release numbers of 500 tags per year provided relatively precise estimates of exploitation rate (hence abundance assuming catch is reported accurately) with CVs of 15% and 11% for the western and eastern stocks, respectively. Although the alternative operating models spanned the widest range of eastern stock migration into the West area, this did not substantially affect the precision of exploitation rates estimates.

The precision of the tagging estimator compares favourably with recent stock assessments where there is additional uncertainty in estimates due to alternative plausible interpretations of data and model assumptions. For example, the base 2017 VPA stock assessment of eastern bluefin tuna estimated exploitation rate in 2015 with a CV of approximately 16% (80% confidence interval of [0.065, 0.096], ICCAT 2017). Running this base assessment but removing a single relative abundance at a time, spanned an additional uncertainty in recent exploitation rate of around 12% (ICCAT 2017). Additional sensitivity analyses -the impact of alternative plausible model configuration, parameterization and data- added a further 18% CV to estimates of current exploitation rate.

For the 2021 western bluefin tuna statistical catch-at-age stock assessment, the base model estimated current exploitation rates with very high precision of approximately 5% (CV). However, simply starting the estimation at varying initial values (a 'jitter analysis'), led to maximum likelihood estimates of spawning abundance in 2020 that varied with a CV of approximately 15% (ICCAT 2021). Running the base assessment but removing a single relative abundance at a time, spanned an additional uncertainty in recent exploitation rate of around 10% (ICCAT 2021). Unlike the current stock assessment approach, gene tagging could provide estimates of management interest that does not rely on a complex interpretation of various, often conflicting data.

These analyses suggest that conventional gene tagging could inform management procedures for the East and West Atlantic management areas that provide expected management performance that is not substantially worse than idealized 'perfect information' constant exploitation rate management. While expected performance was generally very good, the effective number of releases, that is determined by the product of the number of tag releases and the genotyping rate at capture, has important implications for lower tail, 'worst case' management outcomes. It follows that the appropriate investment in such a tagging program should be determined by acceptable levels of risk in terms of yield and biological conservation (among other considerations). This is particularly important as management procedures that fished more strongly at closer to idealized MSY exploitation rate (i.e. ending up with biomass close to BMSY given perfect information) led to declining biomass in later projection years for management procedures using gene tagging data. This further emphasizes the importance of closed-loop simulation testing as a tool in management planning that goes beyond an evaluation of merely the accuracy and precision of tagging estimators.

It is clear from this analysis that the extent of investment in tagging depends on management objectives relating to both long term yields, lower tail biomass outcomes and stock trajectories. Where objectives are less stringent, the required tagging programs may be considerably less intensive. Closed loop MSE-type analyses of the type presented here are required since the exploitation rate of the management procedure determines the precision of the estimates in future years and hence these feedbacks should be accounted for.

Some aspects of these simulations could serve to overestimate the potential performance of gene tagging estimators and related management procedures. On the other hand, performance could be improved with relatively small adjustments that were outside the scope of this research. Natural survival can be estimated by the multiyear brownie model as the number of release years increases. Natural survival was not estimated in this analysis because misspecification of point value for natural survival only serves to proportionally affect the accuracy of exploitation rate estimates. For example, a 20% bias in natural survival leads to a 20% reduction in exploitation rate estimates. Since the focus of this research was management procedure performance, consistent biases in exploitation rate estimation are ignorable because management procedures are generally tuned to obtain comparable biological or yield outcomes. Given the example above, the target exploitation rate of the management procedure using the biased exploitation rate estimator would simply be 20% lower to achieve the same management performance as a management procedure using an unbiased exploitation rate estimator.

The simulation exercise assumed that stock of origin could be identified accurately for all genotyped fish and that there was no genotyping error. Error in stock of origin would not be expected to impact exploitation rate estimation unless one stock had a much higher mark rate and/or exploitation rate, a condition that can be checked for in the simulations and in practice. Similarly to bias in the assumed rate of natural survival, genotyping error would be expected to lead to proportional downward bias in estimates of exploitation rate, which, as explained above, is

ignorable in the context of management procedures that are tuned to performance outcomes. The analyses worked with an idealized model of recapture where the genotyping rate was constant across fleets. If for example, the recapture program were implemented in Japanese markets, it would be necessary to quantify the fraction of fish exported to the Japanese market by fleet, information that was not available in this analysis.

There may be substantial improvements in the estimation performance of the tagging model and the management performance of the gene tagging management procedures given further investigation of more optimal tag release distributions, more advanced tagging models and more sophisticated management procedure algorithms. This analysis adhered to release distributions that have been previously implemented; ultimately all analyses used the historical seasonal/spatial and fleet distribution of electronic tags as the release distribution. There is evidence that this is not ideally balanced with respect to the two stocks requiring a substantially larger number of tags released to obtain comparable precision in western estimates. Further investigation of better performing release distributions might further improve the expected cost-effectiveness of a gene tagging program.

The multiyear Brownie estimator was computationally efficient but was not age-structured and did not account for incomplete mixing of tags that could be related to age-based movement (the operating models assume higher stock viscosity for smaller fish that may be more likely candidates for releases). Age-structured tagging estimators might better characterize these effects and provide improved estimation performance. Age-based estimates of exploitation rate could also be incorporated into management procedures to predict incoming cohort strength, reducing lags in current estimates and improving the reliability of TAC advice. The gene tagging management procedures of these analyses were deliberately very simple, adjusting TACs based on a target exploitation rate relative to the average of the two most recent exploitation rate estimates from the tagging model. More sophisticated management procedure algorithms could account for slope in exploitation rate and include an adjustable exploitation rate target.

At first glance, total tagging program costs of $2.5-4M per year may sound prohibitively high. However, the large majority of those costs relate to the recapture program which could be shared among bluefin stocks originating in the Southern and Pacific Oceans. The Atlantic bluefin fishery is estimated to have an end value of around $1Bn per year (the Southern and Pacific bluefin are valued at approximately $600M and $800M, Pew 2020). It follows that for a very small percentage of the overall end value of the fishery (less than half a percent) it may be possible to implement robust gene tagging management procedures that are fishery-independent, eventually provide independent estimates of natural survival, and could also anchor stock assessments with reliable priors for exploitation rate and abundance. Furthermore, such a tagging program could offer rigorous seafood traceability and an opportunity to coordinate real-time exploitation rate estimation at a centralized location (for example in Japan).

An evaluation of the relative performance of the gene tagging management procedures should be phrased in terms of the cost and expected management performance of the status quo stock-assessment approach for management. The costs of the current system of data collection, processing, assessment (including meetings) is not readily available, but it is likely to be of a similar order of magnitude, while the technical process of stock assessment continues to struggle with large uncertainty in stock productivity and natural survival, and conflicts among data (e.g., Maunder et al. 2021). An extension to this work would be to attempt to evaluate the performance of the stock assessment approach in closed-loop. However, this would require formalizing the assessment and TAC setting process in reproducible code when in reality this is subject to numerous subjective decisions regarding model formulation, parameterization and the interpretation of scientific advice by managers.

The complexity and non-intuitive properties of the current stock assessment paradigm make it accessible to only a relatively narrow range of scientists and stakeholders. Stock assessment requires considerable expertise in implementation and peer review, a large quantity of data and data processing, a number of model assumptions, and is often sensitive to alternative plausible values for central parameters that are not well known (e.g. stock resilience). Together these aspects may not promote confidence in the current stock assessment paradigm as a robust scientific basis for management.

Comparatively, management procedures using tagging data are relatively straightforward to understand and can follow simple rules for the provision of updated TAC advice. This is a principal motivation behind the current MSE framework for Atlantic bluefin tuna and the development of management procedures for bluefin tuna that set TAC advice using relative abundance indices (both fishery independent surveys and catch rate-based fishery dependent indices). Although much simpler than assessments, index-based management procedures must still navigate potential conflicts among indices and require that all indices continue to be collected in the future. If an index-based management procedure for Atlantic bluefin tuna is adopted, an extension to this work could comparatively evaluate the expected performance of the gene tagging management procedures based on established performance metrics.

An evaluation of 'optimal' release numbers and genotyping rate was carried out to determine the least expensive way to achieve a desired precision in the tagging estimator. Because the assumed cost of releasing a tag was almost two orders of magnitude larger than the assumed cost of genotyping, the least expensive tagging design (release numbers and genotyping rate) was essentially invariant to the number of fish caught, the release cost and the genotyping cost. Only the desired level of precision and the exploitation rate of the fishery strongly impact the 'optimal' tagging design. This suggests that a relatively constant tagging design might be successful given that the desired level of precision is likely to be constant and the management procedures aim to achieve a constant target exploitation rate. Rather than focusing on estimation performance, an extension to this work could develop a model that can predict the management performance of a gene tagging management procedure (e.g., annual yields) given release numbers and genotyping rate, and thereby allow for a less arbitrary calculation of 'optimal' tagging designs.

This study confirms that conventional gene tagging may offer a promising basis for calculating management advice for Atlantic bluefin tuna that is cheaper, simpler and more robust than the conventional stock assessment paradigm. Unlike management procedures that make use of relative abundance indices, those that use gene-tagging rely on a single fishery-independent data source and therefore are not susceptible to conflicting data or a discontinuation of a required index. Genetic tagging programs offer a host of other benefits including seafood traceability, quantification of natural survival and the ability to inform stock assessments where status determination is a priority.

Acknowledgments

This research was funded by the FY21 Bluefin Tuna Research Program of the U.S. National Oceanic and Atmospheric Administration (NOAA) (award number: NA21NMF4720304) and supported by The Ocean Foundation (TOF). This draft has benefitted from the guidance and input of Ann Preece (CSIRO), Shana Miller (TOF) and Matt Lauretta (NOAA).

References

Alemany F., Quintanilla L., Velez-Belchí P., García A., Cortés D., Rodríguez J.M., et al. 2010. Characterization of the spawning habitat of Atlantic bluefin tuna and related species in the Balearic Sea (western Mediterranean). Prog. Oceanogr. 86(1–2): 21–38.

Anon. 2021a. Report of the 2021 ICCAT GBYP workshop on close-kin mark recapture for eastern Atlantic bluefin tuna. Collect. Vol. Sci. Pap. ICCAT, 78(3): 198-211.

Anon. 2021b. Report of the 2021 ICCAT GBYP workshop on electronic tagging for Atlantic bluefin tuna. Collect. Vol. Sci. Pap. ICCAT, 78(3): 212-229.

Anon. 2022. Atlantic Bluefin Tuna MSE. International Commission for the Conservation of Atlantic tunas. Retrieved from https://iccat.github.io/abft-mse/ [August 2022]

Andreou, D., Vacquie-Garcia, J., Chucherousset, J. et al. 2011. Individual genetic tagging for teleosts: An empirical validation and guideline for ecologists. Journal of Fish Biology. 80(1): 181-94.

Block B.A., Teo S.L.H., Walli A., Boustany A., Stokesbury M.J.W., Farwell C.J., et al. 2005. Electronic tagging and population structure of Atlantic bluefin tuna. Nature, 434(7037): 1121–1127.

Bradford, R.W., Hill, P., Davies, C,. Grewe, P. 2015. A new tool in the toolbox for large-scale, high throughput fisheries mark-recapture studies using genetic identification. Marine and Freshwater Research http://dx.doi.org/10.1071/MF14423

Bravington, M., Grewe, P., Davies, C. 2014. Fishery-independent estimate of spawning biomass of southern bluefin tuna through identification of close-kin using genetic markers. FRDC Report 2007/034,Hobart, Tas., Australia.

Bravington, M.V., Skaug, H.J., Anderson, E.C. 2016. Close-Kin Mark-Recapture. Statistical Science. 31(2): 259-274. doi: 10.1214/16-STS552.

Butterworth, D.S., Punt, A.E. 1999 Experiences in the evaluation and implementation of management procedures. ICES Journal of Marine Science, 5: 985-998, http://dx.doi.org/10.1006/jmsc.1999.0532.

Carruthers, T.R. 2020. Reference Set Operating models for Atlantic Bluefin Tuna assuming priors for Area-Specific Scale and western Stock Mixing. SCRS/2020/018. Retrieved from https://www.iccat.int/Documents/CVSP/CV077_2020/n_2/CV077020078.pdf [August 2022]

Carruthers, T.R., Butterworth, D.S. 2017. ABT-MSE: an R package for Atlantic bluefin tuna management strategy evaluation. SCRS/2017/225.

Carruthers, T.R., Kell, L.T., Butterworth, D.S., Maunder, M.N., Geromont, H.F., Walters, C., McAllister, M.K., Hillary, R., Levontin, P., Kitakado, T., Davies, C.R. 2015. Performance review of simple management procedures. ICES Journal of Marine Science. 73 (2) 464-482. https://doi.org/10.1093/icesjms/fsv212

Carruthers, T.R., Kimoto, A., Powers, J., Kell, L., Butterworth, D.S., Lauretta, M.V., Kitakado, T. 2015. Structure and Estimation Framework for Atlantic Bluefin Tuna Operating Models. SCRS/2015/179.

CCSBT. 2019. Report of the Twenty Fourth Meeting of the Scientific Committee. Available at: https://www.ccsbt.org/sites/default/files/userfiles/file/docs_english/meetings/meeting_reports/ccsbt_26/report_of_SC24.pdf

Cochrane, K L., Butterworth, D.S., De Oliveira, J.A.A., Roel, B.A., 1998. Management procedures in a fishery based on highly variable stocks and with conflicting objectives: experiences in the South African pelagic fishery. Rev. Fish. Biol. Fisher. 8, 177-214.

Cormack, R. 1964. Estimates of survival from the sighting of marked animals. Biometrika 51 429–438.

Diaz G.A., Turner S.C. 2007. Size frequency distribution analysis, age composition, and maturity of western bluefin tuna in the Gulf of Mexico from the US (1981–2005) and Japanese (1975–1981) longline fleets. Collect. Vol. Sci. Pap. ICCAT, 60(4): 1160–1170.

Fukuda H., Torisawa S., Sawada Y., and Takagi T. 2010. Ontogenetic changes in schooling behaviour during larval and early juvenile stages of Pacific bluefin tuna Thunnus orientalis. J. Fish Biol. 76(7): 1841–1847.

GBYP 2019. ICCAT GBYP Modelling Approaches (MSE details in Sectioni 3.5). Available at: https://www.iccat.int/GBYP/en/modelling.asp

Grewe, P., McDowell, J., Walter, J., Lauretta, M., Gosslien, T., Bravington, M.V., Porch, C., Davies, C.R. 2018. Genomic tools demonstrate excellent potential for estimating a census estimate for Atlantic Bluefin Tuna spawning in the Gulf of Mexico. ICES ASC. Available at: http://www.ices.dk/sites/pub/ASC2018/Abstracts/Forms/DispForm.aspx?ID=601

Galuardi B. and Lutcavage M.E. 2012. Dispersal routes and habitat utilization of juvenile Atlantic bluefin tuna, Thunnus thynnus, tracked with mini PSAT and archival tags. PLoS ONE, 7(5): e37829.

Hernández, C.M., Richardson, D.E., Rypina, I.I., Chen, K., Marancik, K.E., Shulzitski, K., Llopiz. J.K. 2022. Support for the Slope Sea as a major spawning ground for Atlantic bluefin tuna: evidence from larval abundance, growth rates, and particle-tracking simulations. Canadian Journal of Fisheries and Aquatic Sciences. 79(5): 814-824. https://doi.org/10.1139/cjfas-2020-0444

ICCAT 2017. Report of the 2017 ICCAT Bluefin Stock Assessment Meeting. Available at: https://www.iccat.int/Documents/SCRS/DetRep/BFT_SA_ENG.pdf

ICCAT 2020. ABT-MSE: an R package for Atlantic Bluefin Tuna Management Strategy Evaluation. Available at: https://github.com/ICCAT/abft-mse/tree/master/R_package/ABTMSE

ICCAT 2020b. ABT-MSE Shiny App. Available at: http://142.103.48.20:3838/ABTMSE/

ICCAT. 2021. Report of the 2021 western bluefin stock assessment meeting. Retrieved from https://www.iccat.int/Documents/Meetings/Docs/2021/REPORTS/2021_WBFT_SA_ENG.pdf [August 2022]

Lukacs, P.M., Burnham, K.P. 2005. Review of capture–recapture methods applicable to noninvasive genetic sampling. Mol. Ecol. 14 3909–3919.

Lutcavage M.E., Brill R.W., Skomal G.B., Chase B.C., and Howey P.W. 1999. Results of pop-up satellite tagging of spawning size class fish in the Gulf of Maine: do North Atlantic bluefin tuna spawn in the mid-Atlantic? Can. J. Fish. Aquat. Sci. 56(2): 173–177.

Marcek. B.J., Graves, J.E. 2014. An Estimate of Postrelease Mortality of School-Size Bluefin Tuna in the U.S. Recreational Troll Fishery, North American Journal of Fisheries Management, 34:3, 602-608, DOI: 10.1080/02755947.2014.902411

Maunder, M. 2021. Review of the 2021 West Atlantic bluefin tuna assessment. Collect. Vol. Sci. Pap. ICCAT, 78(3): 1114-1124 (2021)

Pew. 2020. 'Netting Billions 2020: A Global Tuna Valuation'. The Pew Charitable Trusts. Retrieved from https://www.pewtrusts.org/en/research-and-analysis/reports/2020/10/netting-billions-2020-a-global-tuna-valuation [August 2022]

Preece, A., Eveson, P., Davies, C., Grewe, P., Hillary, R., Bravington, M. 2015. Report on gene-tagging design study. Commonwealth Scientific and Industrial Research Organisation (CSIRO). CCSBT-ESC/1509/18. Available at address

Preece, A. L., Eveson, P. J., Grewe, P.M., Bradford, R., Clear, N., Aulich, J., Lansdell, M., Cooper, S., Hartog, J. 2018b. The 2018 SBT gene-tagging program. Commonwealth Scientific and Industrial Research Organisation (CSIRO). CCSBT-OMMP/2006/04. Available at:

https://www.ccsbt.org/en/system/files/OMMP11_05_CCSBT_GT2018Program_DataForMPandOM.pdf [accessed December 2021]

Puncher G.N., Cariani A., Maes G.E., Van Houdt J., Herten K., Cannas R., et al. 2018. Spatial dynamics and mixing of bluefin tuna in the Atlantic Ocean and Mediterranean Sea revealed using next-generation sequencing. Mol. Ecol. Resour. 18(3): 620–638.

Richardson D.E., Marancik K.E., Guyon J.R., Lutcavage M.E., Galuardi B., Lam C.H., et al. 2016. Discovery of a spawning ground reveals diverse migration strategies in Atlantic bluefin tuna (Thunnus thynnus). Proc. Natl. Acad. Sci. U.S.A. 113(12): 3299–3304.

Rooker J.R., Arrizabalaga H., Fraile I., Secor D.H., Dettman D.L., Abid N., et al. 2014. Crossing the line: migratory and homing behaviors of Atlantic bluefin tuna. Mar. Ecol. Prog. Ser. 504: 265–276.

Stokesbury M.J.W., Neilson J.D., Susko E., Cooke S.J. 2011. Estimating mortality of Atlantic bluefin tuna (Thunnus thynnus) in an experimental recreational catch-and-release fishery. Biol Conserv. 144:2684–91.

Table 1. Effective number of annual tag releases (annual release number multiplied by genotyping rate of recaptured fish) required to obtain specified coefficient of variation in exploitation rate estimates from the multiyear brownie model (year 2028). These numbers were calculated using the fitted linear model used to predict simulated mean CVs (Eqn 2, **Figure** 6).

Coefficient of variation of harvest rate estimates	Western Stock										Eastern Stock									
	Exploitation rate										Exploitation rate									
	1%	2%	3%	4%	5%	6%	7%	8%	9%	10%	1%	2%	3%	4%	5%	6%	7%	8%	9%	10%
10.0%	2888	2045	1624	1362	1179	1043	938	854	784	726	2150	1463	1132	930	792	691	613	551	501	459
12.5%	2223	1502	1161	956	816	715	637	575	525	483	1597	1035	779	628	528	455	401	358	323	295
15.0%	1763	1149	871	708	599	520	461	414	376	345	1233	770	569	453	377	323	282	251	226	205
17.5%	1433	908	678	545	458	395	349	312	283	258	981	596	433	342	282	241	209	185	166	151
20.0%	1187	735	542	433	361	311	273	244	220	201	799	474	341	267	219	186	162	143	128	115
22.5%	1000	607	443	352	292	251	219	195	176	160	663	387	275	214	175	148	128	113	101	91
25.0%	853	510	370	291	241	206	180	160	144	131	559	321	227	176	143	121	105	92	82	74
27.5%	737	435	313	246	203	173	151	134	120	109	478	271	190	147	119	101	87	76	68	61
30.0%	643	375	268	210	173	147	128	113	102	92	413	232	162	125	101	85	73	64	57	51
32.5%	566	326	232	181	149	126	110	97	87	79	361	201	139	107	87	73	62	55	49	44
35.0%	502	287	203	158	130	110	95	84	76	69	318	175	121	93	75	63	54	47	42	38

Table 2. Calculations of the least expensive combination of release number (No. Releases) and genotype rate at recapture, calculated from the linear model approximations of precision for the eastern and western stocks (Eqns. 2, 3, **Figure** 6). Coefficient of variation refers to the specified level of precision in exploitation rate estimates after 6 tag release years (Eqn. 2). No. fish caught refers to the total landed catches of bluefin tuna in numbers. Release cost is the cost of releasing a gene tag at sea. Genotyping cost is the cost per sample. The default values of the numerical optimization are highlighted by the gray shaded rows (hence these rows are identical): a target estimator precision (σ) of 20%; a fishery exploitation rate (u) of 4%; 250,000 fish are caught and landed each year (n_{caught}); an in-water tag release costs (C_{rel}) $1000; a genotyping sample at recapture costs (C_{gen}) $15. All dollar values are expressed as US$.

	Precision calculated by western stock linear model						Precision calculated by eastern stock linear model					
	Total Costs						Costs					
Variable	Release ($ 000s)	Capture ($ 000s)	Capture /Release	No. Releases	Genotyping rate	Effective No. Releases	Release ($ 000s)	Capture ($ 000s)	Capture /Release	No. Releases	Genotyping rate	Effective No. Releases
Coefficient of variation (σ)												
0.1	1451	3520	2.43	1451	93.9%	1362	1055	3305	3.13	1055	88.1%	930
0.125	1082	3314	3.06	1082	88.4%	956	745	3161	4.24	745	84.3%	628
0.15	827	3210	3.88	827	85.6%	708	564	3010	5.34	564	80.3%	453
0.175	659	3099	4.70	659	82.6%	545	448	2859	6.38	448	76.3%	342
0.2	543	2987	5.50	543	79.7%	433	369	2712	7.35	369	72.3%	267
0.225	458	2876	6.27	458	76.7%	352	314	2563	8.18	314	68.4%	214
0.25	395	2766	7.00	395	73.8%	291	272	2426	8.92	272	64.7%	176
Exploitation rate (u)												
0.02	855	3224	3.77	855	86.0%	735	586	3034	5.17	586	80.9%	474
0.03	656	3097	4.72	656	82.6%	542	447	2859	6.39	447	76.2%	341
0.04	543	2987	5.50	543	79.7%	433	369	2712	7.35	369	72.3%	267
0.05	468	2891	6.17	468	77.1%	361	319	2578	8.08	319	68.7%	219
0.06	415	2804	6.75	415	74.8%	311	282	2473	8.76	282	65.9%	186
0.07	375	2726	7.26	375	72.7%	273	256	2372	9.28	256	63.3%	162
0.08	345	2644	7.66	345	70.5%	244	233	2292	9.82	233	61.1%	143
No. fish caught (000s) (n_{caught})												
150	541	1797	3.32	541	79.9%	433	367	1636	4.45	367	72.7%	267
175	542	2095	3.87	542	79.8%	433	368	1905	5.18	368	72.6%	267
200	542	2393	4.41	542	79.8%	433	368	2175	5.91	368	72.5%	267
225	543	2690	4.96	543	79.7%	433	369	2444	6.63	369	72.4%	267
250	543	2987	5.50	543	79.7%	433	369	2712	7.35	369	72.3%	267
275	543	3284	6.05	543	79.6%	433	371	2968	8.00	371	71.9%	267
300	543	3581	6.59	543	79.6%	433	372	3233	8.70	372	71.8%	267
Release cost ($ per fish) (C_{rel})												
500	274	2960	10.80	548	78.9%	433	187	2678	14.33	374	71.4%	267
1000	543	2987	5.50	543	79.7%	433	369	2712	7.35	369	72.3%	267
1500	813	2994	3.68	542	79.8%	433	551	2723	4.94	368	72.6%	267
2000	1082	2999	2.77	541	80.0%	433	733	2731	3.73	367	72.8%	267
2500	1351	3002	2.22	540	80.0%	433	914	2737	2.99	366	73.0%	267
3000	1620	3004	1.85	540	80.1%	433	1095	2742	2.50	365	73.1%	267
3500	1889	3006	1.59	540	80.2%	433	1276	2747	2.15	364	73.2%	267
Genotyping cost ($ per processed sample) (C_{gen})												
10	542	1996	3.68	542	79.8%	433	368	1815	4.94	368	72.6%	267
12.5	542	2492	4.59	542	79.7%	433	368	2264	6.15	368	72.5%	267
15	543	2987	5.50	543	79.7%	433	369	2712	7.35	369	72.3%	267
17.5	543	3482	6.41	543	79.6%	433	371	3145	8.47	371	71.9%	267
20	544	3977	7.31	544	79.5%	433	372	3586	9.64	372	71.7%	267
22.5	544	4471	8.22	544	79.5%	433	373	4027	10.80	373	71.6%	267
25	547	4942	9.04	547	79.1%	433	373	4475	12.00	373	71.6%	267

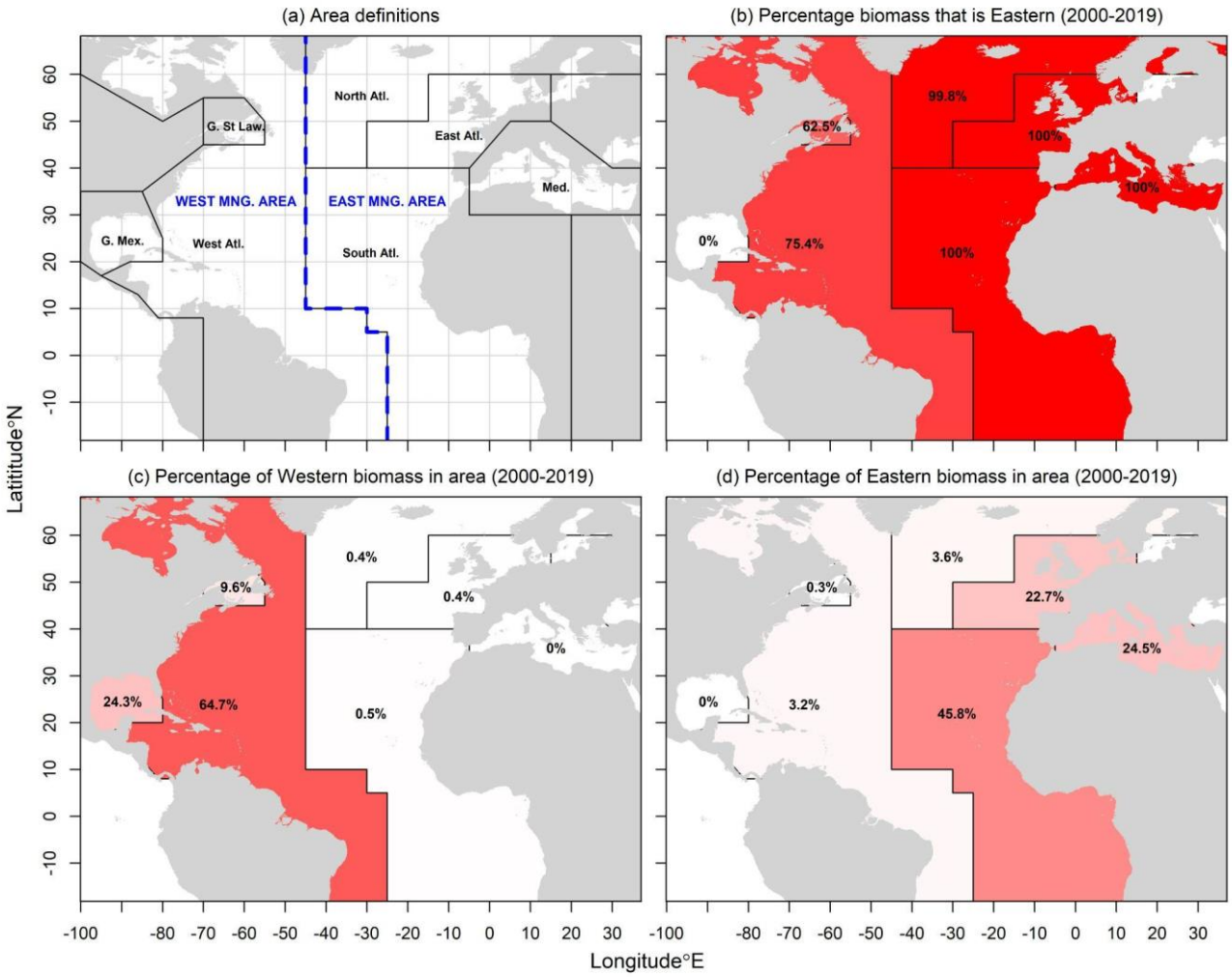

Figure 1. Definition of East and West management areas (blue), the seven areas of the spatial operating model (black) and estimates of recent stock distribution and mixing for the Reference Case 'Low abundance' operating model where darker red areas are those of higher relative biomass.

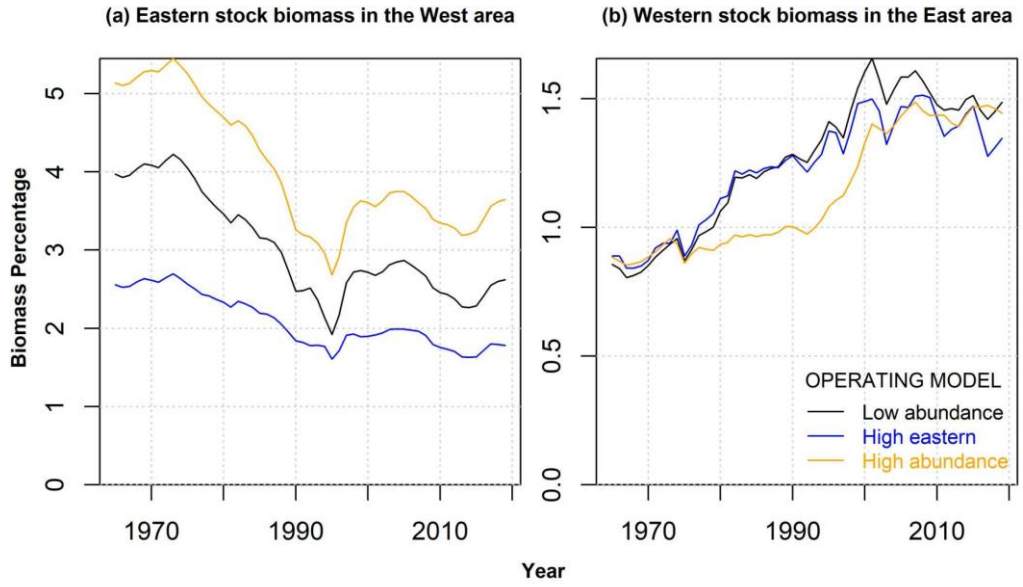

Figure 2. The percentage biomass of each stock in the opposite management area for the Low abundance operating model.

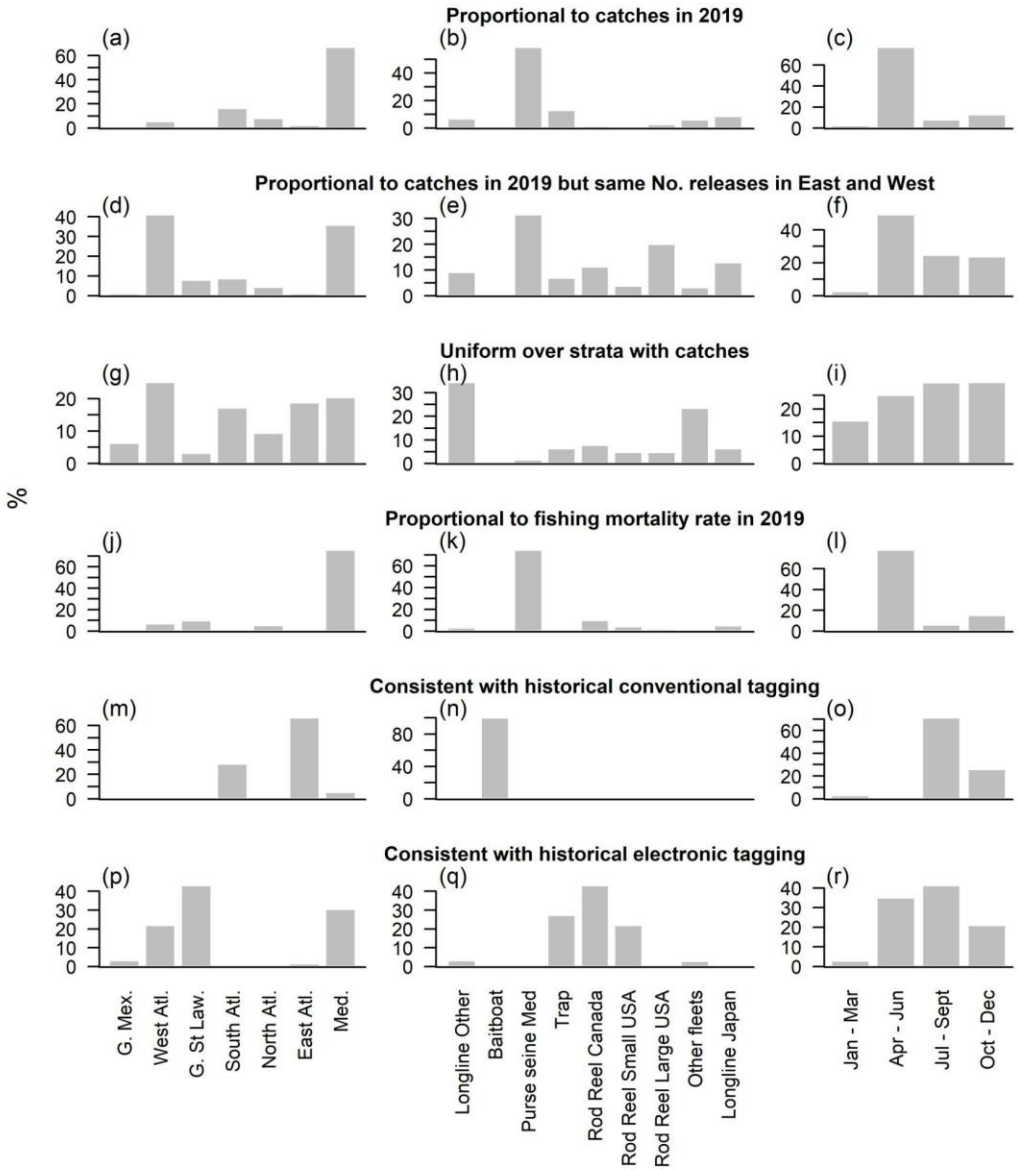

Figure 3. The distribution of tag releases by area, fleet and season for the 6 tag release distributions (row).

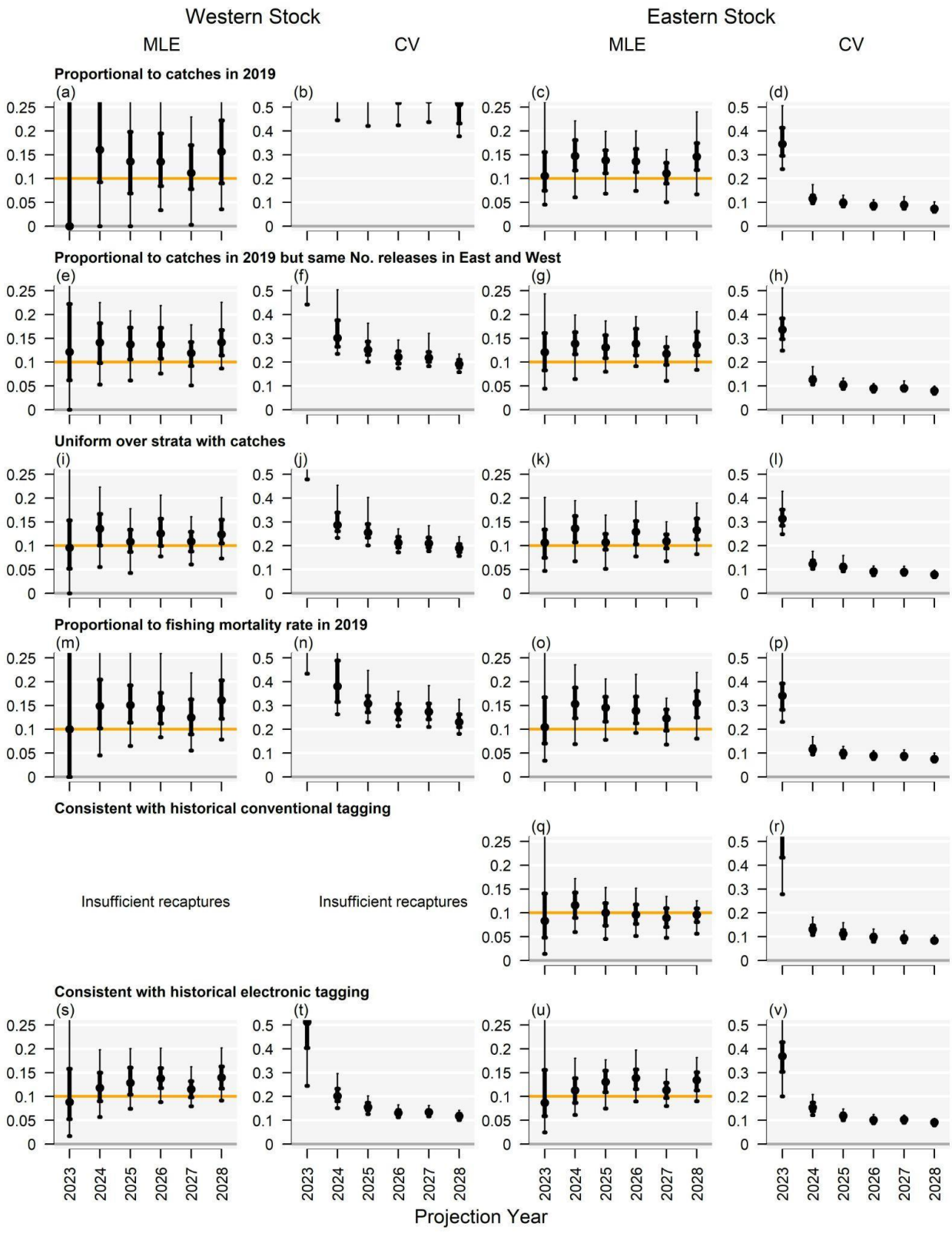

Figure 4. Annual exploitation rate (MLE) and coefficient of variation (CV) estimated by the multiyear brownie model in 2028 for projection years 2024 - 2028 (150 simulations) for the 'Low abundance' operating model. Genetic tagging data were simulated from the individual tagging model subject to six tagging designs (each is represented by a row of figure panels) assuming a 10% exploitation rate (orange horizontal line) and 500 effective releases per year. Plotted points represent the mean values, thick and thin bars represent the interquartile and 90% interquantile ranges, respectively.

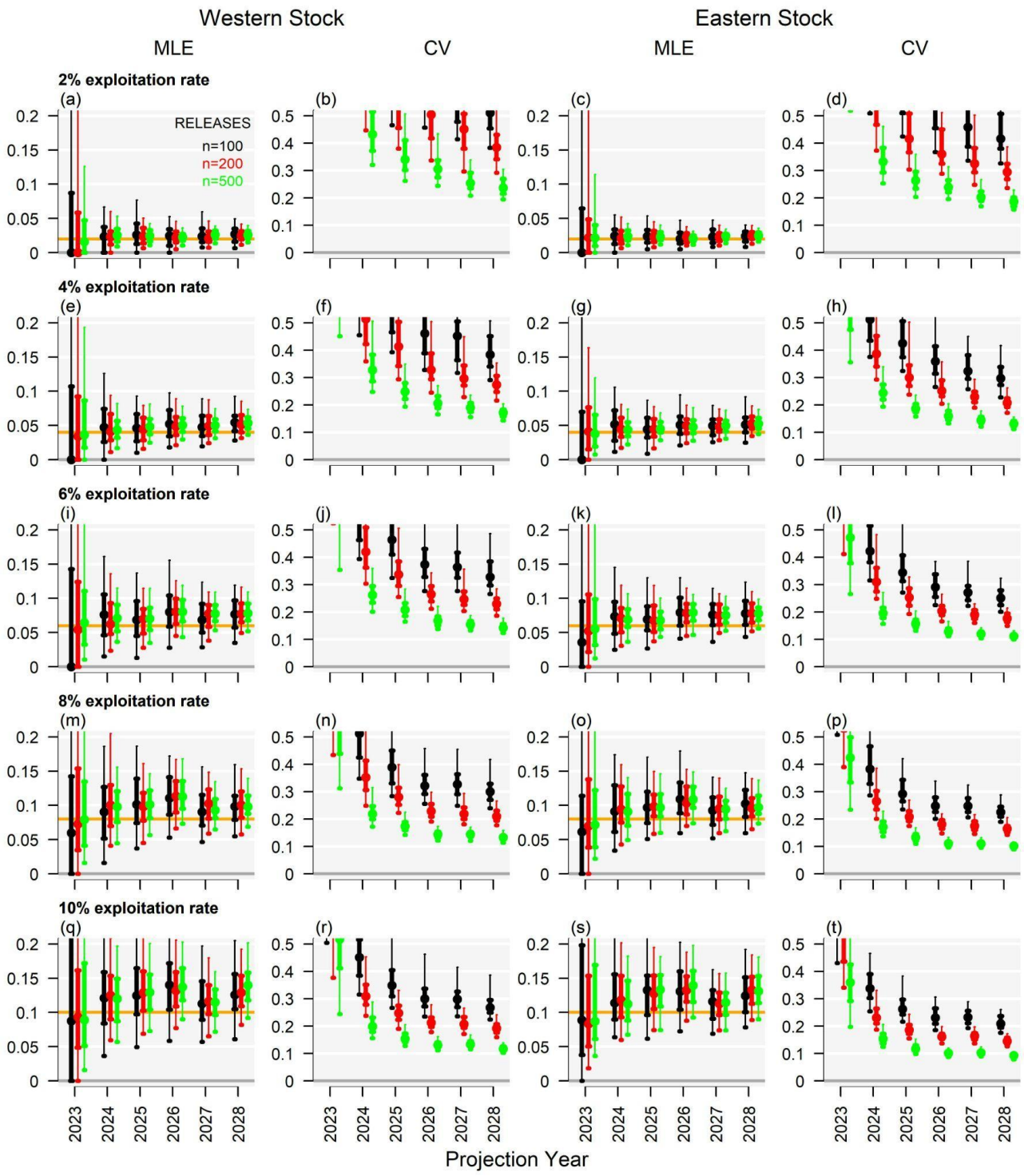

Figure 5. As **Figure 4** but estimation performance is presented across varying levels of exploitation rate (row) and release number (color), given a tagging design 'consistent with historical electronic tagging' and the 'Low abundance' operating model.

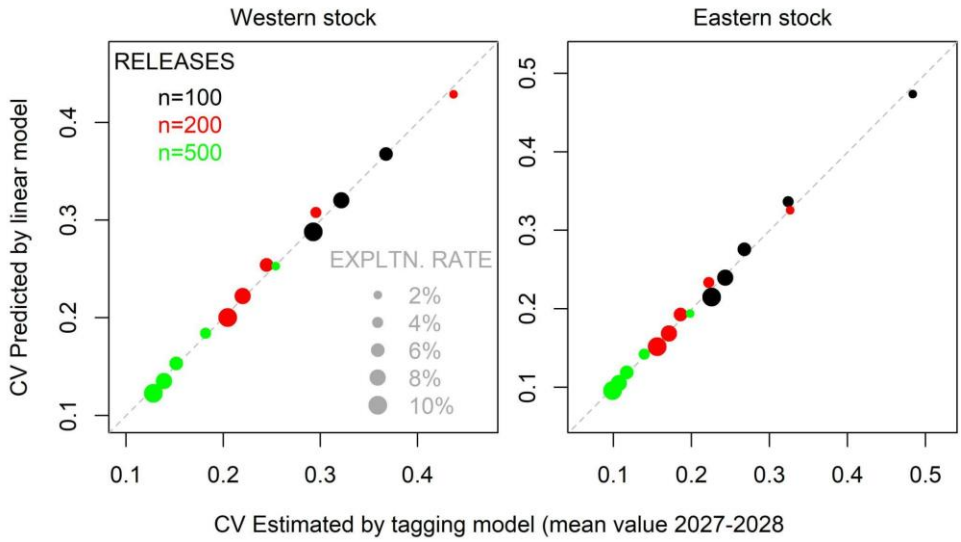

Figure 6. Linear model predictions of the precision in exploitation rate (coefficient of variation, CV) estimated by the multiyear brownie model in the year 2028 under varying release numbers and exploitation rates (releases following the distribution of historical electronic tagging releases)(Eqn. 2).

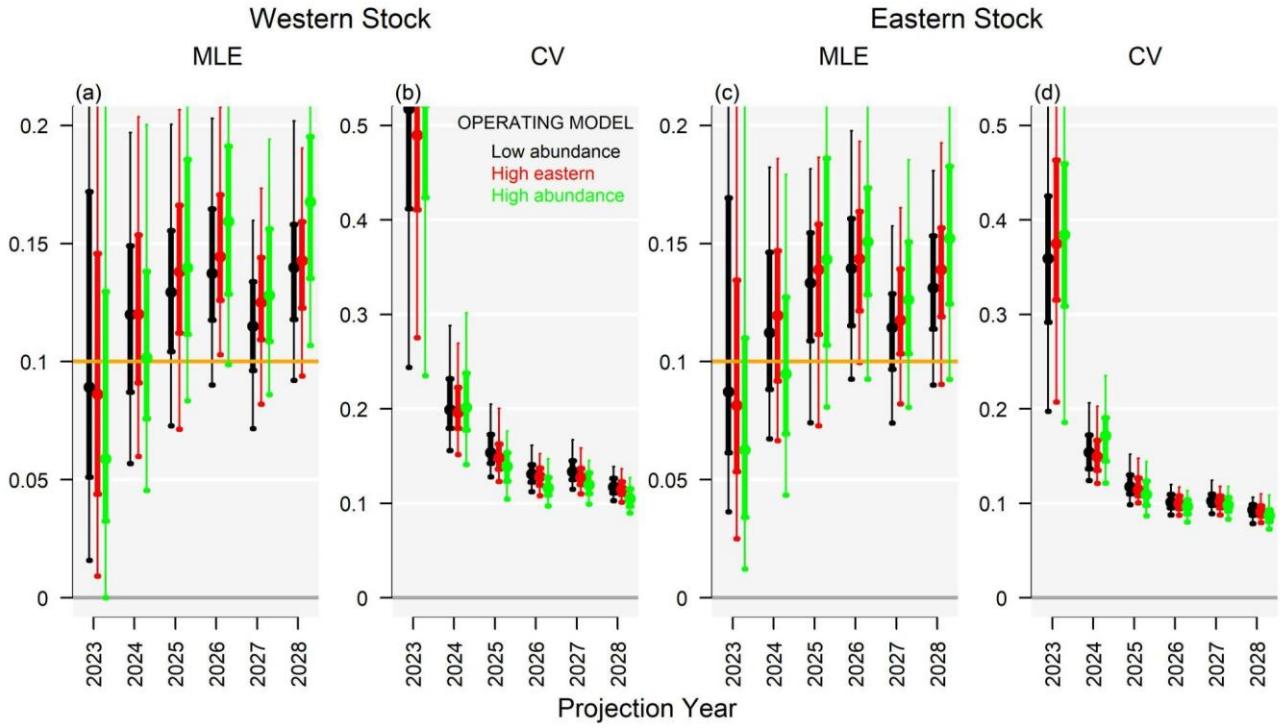

Figure 7. As **Figure 4**, but showing the effect of operating model dynamics on the bias and precision of estimates of exploitation rate calculated by the multiyear brownie model given a 10% exploitation rate, 500 effective annual releases of tags, and releases distributed according to historical electronic tagging releases.

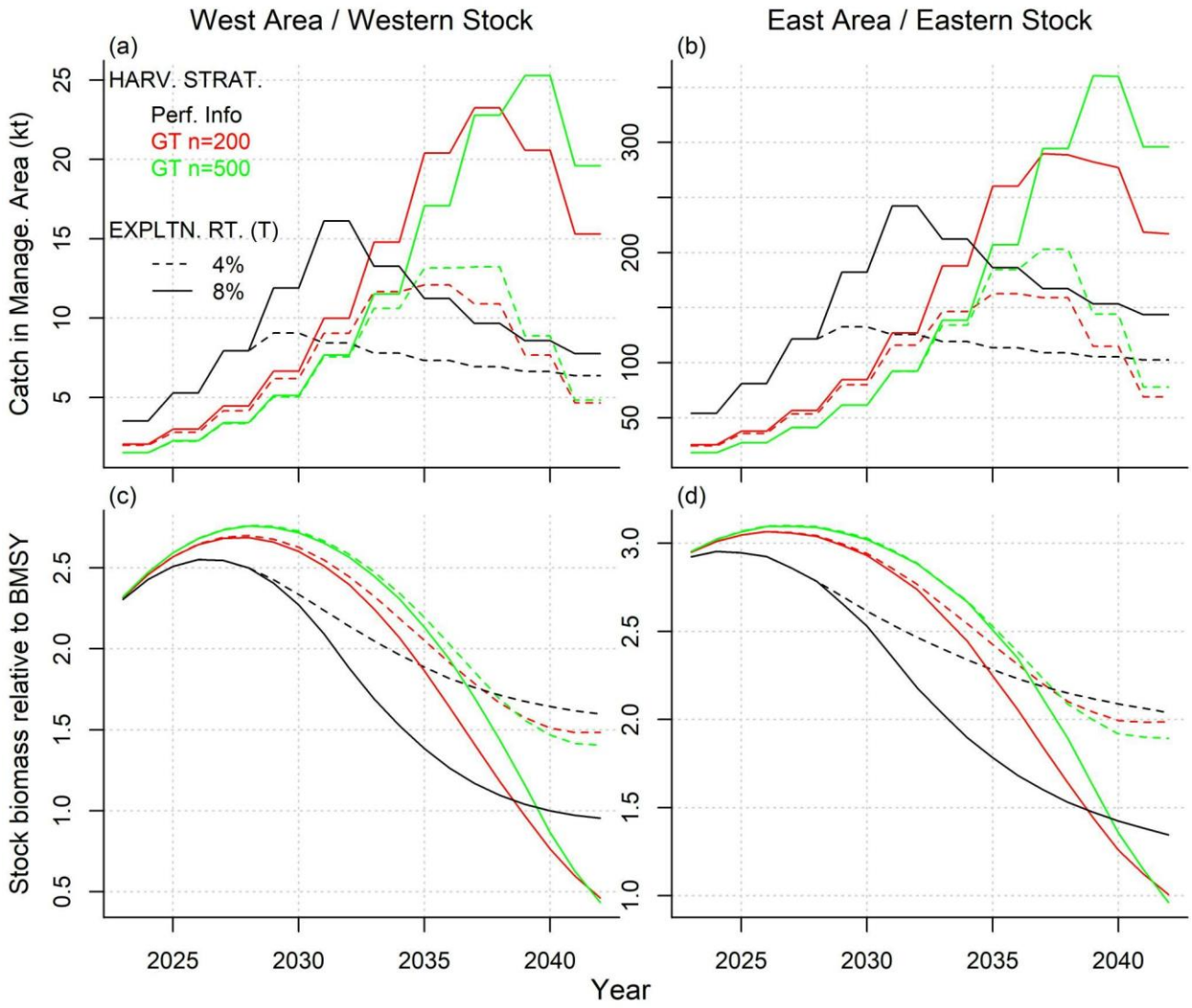

Figure 8. Mean projected annual catches and spawning stock biomass relative to BMSY for the gene tagging and perfect information management procedures (harvest strategies) at target exploitation rates of 4% and 8% (150 simulations) for the Reference Case 'Low abundance' operating model and releases distributed according to historical electronic tagging releases.

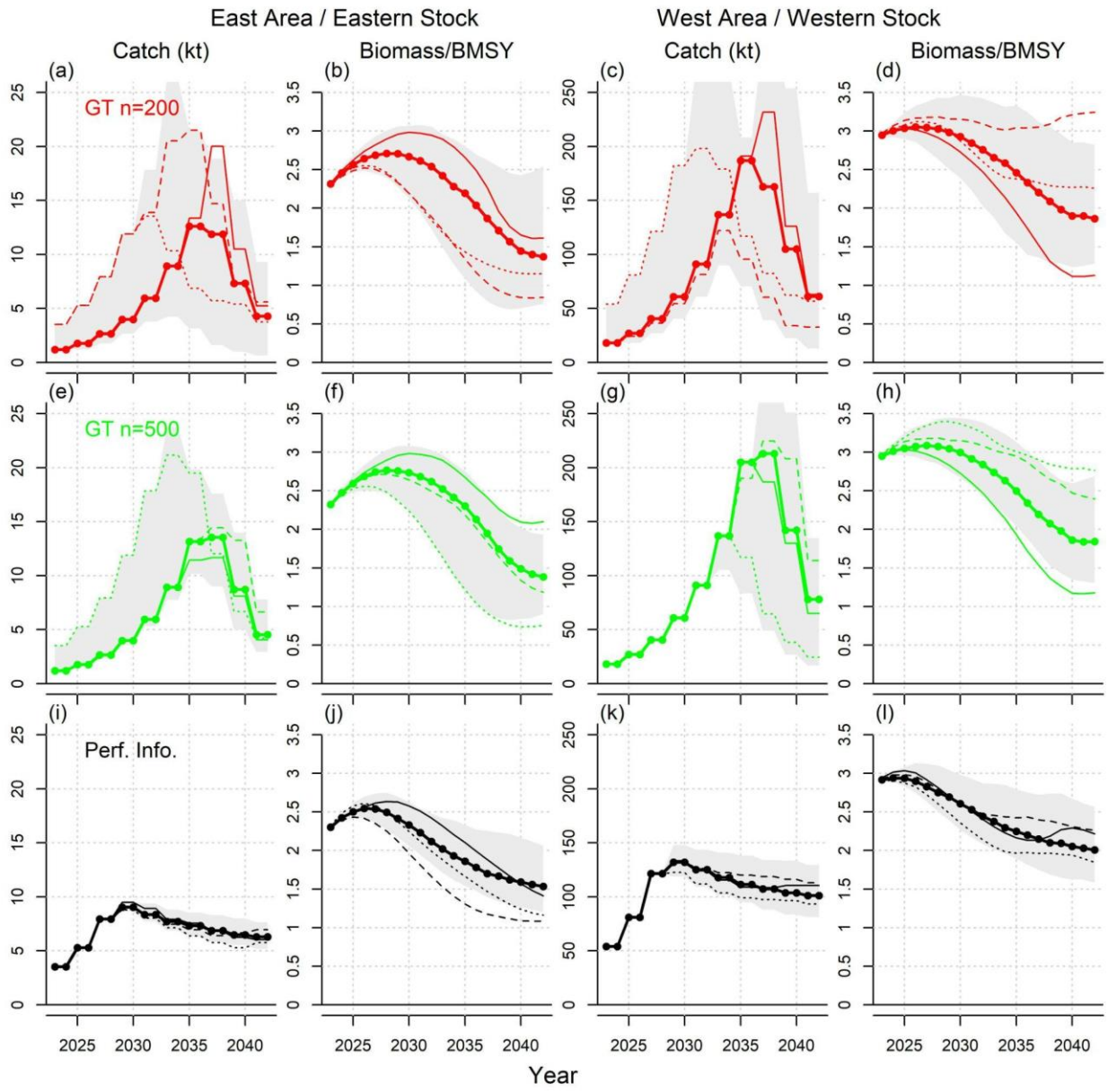

Figure 9. Projections of catch and biomass relative to BMSY for the Reference Case 'Low abundance' operating model and releases distributed according to historical electronic tagging releases, for management procedures with a target exploitation rate of 4%. The thick line with points represents the mean value, the gray shaded area represents the 90% interquantile range (150 simulations), and the three thin lines are three individual simulations.

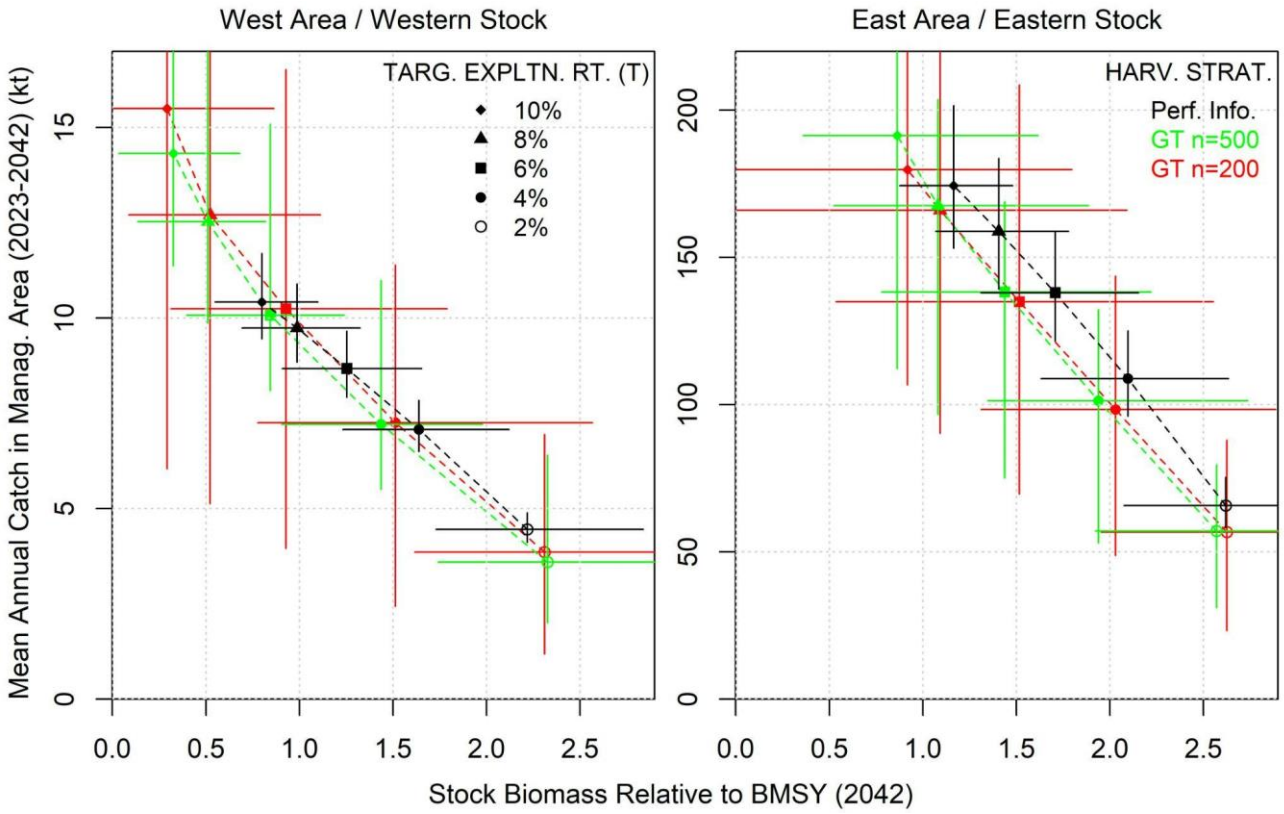

Figure 10. Biomass and yield performance of management procedures (harvest strategies) aiming to fish at constant exploitation rates from 2% to 10%. Results are presented for the 'Low abundance' operating model with releases distributed according to historical electronic tags. Management procedures are paired and applied in both management areas simultaneously (i.e. the 2% management strategy in the West area was run in tandem with the 2% management procedure in the East area). A 'perfect information' management procedure calculated the TAC from the specified exploitation rate given perfect knowledge of current vulnerable biomass in each management area. Management procedures using the gene tagging estimator (GT) were also evaluated given two numbers of effective annual tag releases (n=200 and n=500 per year). Biological management performance is presented as stock biomass at the end of the 20 year projection (2042) and yield management performance as the average annual catch over the 20-year projection (2023-2042). Points represent the mean value and bars the 90% interquantile range (150 simulations).

Appendix A

Individual Tagging Model

App. Table. 1. Subscript terms for fishery dynamics model and individual tagging model

Subscript	Description
t	Time step (year – quarter)
a	Age
r	Area (area moving to)
k	Area moving from
f	Fleet
s	Stock
i	Tag number

App. Table 2. Parameter and variable names for fishery dynamics model and individual tagging model. Note that the MSE calculations include have *F*, *N*, *Y* and *R* parameters that vary by simulation. For simplicity a subscript for simulation is not included in the mathematical description of the model below.

Parameter / Variable	Description
$F_{t,a,r,f}$	Instantaneous fishing mortality rate
$M_{t,a,s}$	Instantaneous natural mortality rate (m^{-1})
$N_{t,a,r,s}$	Number of individual fish at liberty (both untagged and tagged)
$R_{t,r}$	Released tags
$T_{i,t,a,r,s}$	Tagged fish at liberty
$H_{i,t,s}$	Observed tag history
$P_{t,a,r,s}$	Expected magnitude of exploitation
$S_{i,t,a,r,s}$	Individual survival switch
$Y_{t,a,r,k,s}$	Movement probability (from area *k* to area *r*)
$U_{t,a,r,s}$	Harvest rate (fraction of individual caught by fishing)
$L_{i,t,a,s}$	Probability of recapturing a tag
$K_{i,t,a,s}$	Recapture switch
$W_{i,t,a,s}$	Probability of harvesting a tag

App. Table 3. Equations for individual tagging model in order of operation: survival, movement, release / recapture, harvesting.

Tagged fish survival	
Individual tag survival *S*, is sampled from a binomial distribution: $$S_{i,t,a,r,s} \sim bin(n = 1, p = e^{-M_{t,a,s}})$$ If the tagged fish survives, the initial tagged state $\dot{T}$ is unchanged from the final state T of the previous time step. Otherwise the tagged fish dies (S = 0) and the tagged state becomes zero: $$\dot{T}_{i,t,a,r,s} = S_{i,t,a,r,s} T_{i,t-1,a,r,s}$$	1 2
Tagged fish movement	
The tagged fish are moved according to a seasonal, spatial, age-specific, stock-specific movement matrix *Y* created a distribution of expected tag fractions $\ddot{T}$, among areas *r*: $$\ddot{T}_{i,t,a,r,s} = \sum_{k} \dot{T}_{i,t-1,a,r,s} Y_{t-1,a,r,k,s}$$	3
Tag releases	
Given a defined number of tag releases *R*, in each time step *t*, and area *r*, the tag is assigned to age *a*, and stock *s*, according to a multinomial distribution: $$T_{i,t,a,r,s} \sim multinom\left(n = R_{t,r}, \quad p = \frac{P_{t,a,r,s}}{\sum_a \sum_s P_{t,a,r,s}}\right)$$ where the probability of a tag release in a given age and stock is determined by the relative frequency of numbers *N*, multiplied by fishing mortality rate *F*: $$P_{t,a,r,s} = N_{t,a,r,s} \sum_{f} F_{t,a,r,f}$$	4 5
Initializing tag history	
The observed tag history *H*, is recorded for each tag *i*, and uses integer numbers to record a release (*H* = 1), no new information (*H* = 0), an at-sea recapture and re-release (*H* > 1 where the number reflects the number of at-sea recaptures of the same tagged fish) or a negative integer (H < 0) that reflects the index of the fleet that harvested the tagged fish. When a tag is initially released *H* is updated to reflect a new release: $$H_{i,t,s} = \begin{cases} 1 & \sum_r T_{i,t,a,r,s} = 1 \\ 0 & \sum_r T_{i,t,a,r,s} = 0 \end{cases}$$	6
Tag recaptures	
Tag recapture probability is determined by the total harvest rate *U*, calculated from the fishing mortality rate *F*, across all fleets f: $$U_{t,a,r} = \left(1 - e^{-\sum_f F_{t,a,r,f}}\right)$$ The recapture probability *L* of a given tag *i*, in a given area *r*, is the total harvest rate multiplied by the proportion of the population numbers that is represented by the tag:	7

$$L_{i,t,a,r,s} = U_{t,a,r}\frac{\ddot{T}_{i,t,a,r,s}}{N_{t,a,r,s}} \quad (8)$$

Tag recapture events are sampled from a binomial distribution determined by the product of recaptures probabilities over all areas:

$$K_{i,t,a,s} \sim bin\left(1, 1 - \prod_r 1 - L_{i,t,a,r,s}\right) \quad (9)$$

Given a tag recapture (K = 1) the recapture is assigned to a single area based on the probability of recapture by area:

$$T_{i,t,a,r,s} = \begin{cases} \sim multinom\left(1, \frac{L_{i,t,a,r,s}}{\sum_r L_{i,t,a,r,s}}\right) & K_{i,t,a,s} = 1 \\ \ddot{T}_{i,t,a,r,s} & K_{i,t,a,s} = 0 \end{cases} \quad (10)$$

The tag history is updated to log a new recapture of the same tag:

$$H_{i,t,s} = \begin{cases} H_{i,t-1,s} + 1 & K_{i,t,a,s} = 1 \\ H_{i,t-1,s} & K_{i,t,a,s} = 0 \end{cases} \quad (11)$$

Tag harvesting

Tag harvesting events are sampled from a binomial distribution determined by the product of the harvest rate over all areas:

$$W_{i,t,a,s} \sim bin\left(1, 1 - \prod_r 1 - U_{t,a,r}\right) \quad (12)$$

Given a harvesting event, the area the tag was recaptured in is determined by a multinomial distribution where the area is sampled based on the relative magnitude of the harvest rate in each area:

$$T_{i,t,a,r,s} = \begin{cases} \sim multinom\left(1, \frac{U_{t,a,r}}{\sum_r U_{t,a,r}}\right) & W_{i,t,a,s} = 1 \\ 0 & W_{i,t,a,s} = 0 \end{cases} \quad (13)$$

Similarly, the fleet responsible for recapture is recorded in the tag history as the negative value of the fleet index (an integer value). The fleet index is sampled from a multinomial distribution based on the relative magnitude of fishing mortality rate F:

$$H_{i,t,s} = \begin{cases} -\sim multinom\left(1, \frac{F_{t,a,r,f}}{\sum_f F_{t,a,r,f}}\right) & W_{i,t,a,s} = 1 \\ H_{i,t-1,s} & K_{i,t,a,s} = 0 \end{cases} \quad (14)$$

Appendix B

Multiyear Brownie Estimator

The Brownie model (Brownie 1978, 1985; as cited in Hoenig et al. 1998) is an approach for estimating annual survival in a population from multiple years of tagged releases of animals and subsequent recaptures over time. The model has proven to be flexible framework and allows for modifications to relax strict assumptions, such as the immediate, complete mixing of the tags into the population and complete tag retention (see Hoenig et al. 1998 and Waterhouse and Hoenig, 2011, as examples), typically associated with tagging models. Methodological improvements have re-parameterized of survival into fishing mortality and natural mortality components, increasing its utility for fisheries assessment (Hoenig et al. 1998).

In a Brownie model, we first consider the number of tags released in year *i* ($R_{i,s}$) for stock *s* (for simplicity, subscript *s* is dropped hereafter but the Brownie estimator is applied independently for each stock). Assuming immediate mixing of tagged fished into the population, the number of tagged fish $T_{i,j}$ of cohort *i* in year *j* would be

$$T_{i,j} = \begin{cases} R_i & i = j \\ T_{i,j-1} \exp\left(-\left[F_{j-1} + M\right]\right)\phi & i < j \end{cases}$$

where F_j is the fishing mortality in year *j*, M is the natural mortality, and ϕ is the chronic tag retention rate (here assumed to be one).

For Atlantic bluefin tuna, a modification is made for incomplete mixing within the year of release:

$$T_{i,j} = \begin{cases} R_i & i = j \\ T_{i,j-1} \exp\left(-\left[q_{j-i-1}F_{j-1} + M\right]\right)\phi & i < j \end{cases}$$

The catchability of within-year recaptures, relative to those in later years, is expected to be higher as fish remain in the area in which they were tagged. This difference is represented by parameter q_k where $k = j - i$ is the time lag between release and recapture in calendar years and is modeled as a separable effect on F_j. After one year, tags can be considered to be fully mixed into the population, i.e., $q_k = 1$ for $k = 1, 2, 3, \ldots$

For years $i \leq j$, the recaptures $H_{i,j}$ is predicted by the Baranov equation,

$$H_{i,j} = \frac{q_{j-i}F_j}{q_{j-i}F_j + M}\left(1 - \exp\{-[q_{j-i}F_j + M]\}\right)T_{i,j}\lambda$$

where λ is the tag reporting rate. Here, a 100% reporting rate was assumed (λ = 1).

The log-likelihood L of the model uses a multinomial distribution for the fates of tag cohort *i* over years $i \leq j$,

$$L = \sum_i \left(\sum_{i \leq j} \left[h_{i,j} \log\left(\hat{p}_{i,j}\right)\right] + \left(R_i - \Sigma_i h_{i,j}\right) \log\left(1 - \Sigma_i \hat{p}_{i,j}\right) \right)$$

where $h_{i,j}$ is the observed number of recaptures, $\hat{p}_{i,j} = \hat{H}_{i,j}/R_i$ is the predicted proportion of tags recaptured, with the hat operator (^) denoting an estimate, and the last term in the equation is the likelihood component of tags that have not been seen since release. The estimated parameters were F_j and $q_{k=0}$.

The Brownie model was implemented in Template Model Builder (TMB), an R package for implementing rapid, complex models (Kristensen et al. 2016).

References

Brownie, C., Anderson, D.R., Burnham, K.P., and Robson, D.S. 1978. Statistical inference from band recovery data - a handbook. U.S. Fish Wildl. Serv. Resour. Publ. No. 131.

Brownie, C., Anderson, D.R., Burnham, K.P., and Robson, D.S. 1985. Statistical inference from band recovery data: a handbook. 2nd ed. U.S. Fish Wildl. Serv. Resour. Publ. No. 156.

Hoenig, J.M., Barrowman, N.J., Pollock, K.H., Brooks, E.N., Hearn, W.S., and Polacheck T. 1998. Models for tagging data that allow for incomplete mixing of newly tagged animals Can. J. Fish. Aquat. Sci. 55: 1477-1483.

Kristensen, K., Nielsen, A., Berg, C.W., Skaug, H., and Bell, B.M. 2016. TMB: Automatic Differentiation and Laplace Approximation. J. Stat. Soft. 70. doi:10.18637/jss.v070.i05

Waterhouse, L., and Hoenig, J.M. 2011. Instantaneous-Rates Tagging Models Allowing for Delayed Mixing of Newly Tagged Cohorts: Partial Year Tabulation of Recaptures. N. Am. J. Fish. Manage. 31: 995-1004.