SPECIFICATIONS FOR MSE TRIALS FOR ATLANTIC BLUEFIN TUNA
Version 22-05: November 17 2022

CONTENTS

1 BASIC CONCEPTS AND STOCK STRUCTURE

This first item intends to cover only the broadest overview issues. More detailed technical specifications are included under subsequent items (**Table 1.1**).

Table 1.1. Terminology used in this document.

Term	Description
Candidate Management Procedure (CMP)	Some combination of monitoring, assessment, harvest control rule and management action designed to meet the stated objectives of a fishery. A fully specified harvest strategy that has been simulation tested for performance and adequate robustness to uncertainties is often referred to as a Management Procedure.
Operating Model (OM)	A model representing a plausible scenario for stock and fishery dynamics that is used to simulation test the management performance of CMPs.
Reference Grid / Robustness Set	Reference Grid: The operating models that represent the most important uncertainties in stock and fishing dynamics, which are used as the principal basis for evaluating CMP performance. The reference operating models are specified according to factors (e.g., natural mortality rate) that have multiple levels (possible scenarios for each factor, e.g., high / low natural mortality rate). Reference operating models are organized in a usually fully crossed orthogonal 'grid' of all factors and levels (see **Tables 9.1** and **9.2**). Robustness Set: Other potentially important uncertainties in stock and fishing dynamics may be included in a Robustness Set of operating models that provide additional tests of CMP performance robustness. They can be used to further discriminate between CMPs. Compared to the Reference Grid operating models, the Robustness Set models will be typically less plausible and/or influential on performance.
Baseline / Alternative options	All OMs (both Reference Grid and Robustness Set) are specified according to a set of Baseline assumptions. For completeness, alternative options are described in this document which could be considered in the specification of operating models in future iterations of the MSE.
Stratum	One of the seven spatial areas described in **Figure 1.1**.
Area, Management Area	The two management areas are delineated by 45 degrees west in the North Atlantic. The West management area includes areas (or strata) 1-3 of **Figure 1.1**. Areas 4-7 constitute the East management area.
Stock	In this document the term 'Stock' refers to fish originating from a natal spawning region, either in the Mediterranean (Eastern Stock) or in the Gulf of Mexico / Slope Sea (Western Stock). Hence 'Eastern' and 'Western' may be used to describe fish according to their stock of origin, for example 'Eastern fish' are fish that spawn in the Mediterranean. Eastern fish (of the Eastern Stock) can migrate to the West area (management area) and Western fish (of the Western Stock) can migrate to the East area (management area). However, it is assumed that no Eastern fish enter the Gulf of Mexico and no Western fish enter the Mediterranean. 'Stock' in this document should not be confused with 'stock' as defined in previous stock assessments that are based on Management Area.
Season	A quarterly division of the calendar years used in the OMs to account for seasonal migration and stock mixing: (1) Jan-Mar, (2) Apr – Jun, (3) July – Sept and (4) Oct-Dec. The term 'spatio-temporal stratum' refers to one of the seven spatial strata in a particular year and quarter (e.g., stratum 6, quarter 3 in 1981).

Table 1.1. Continued.

Term	Description
Terminal year	The most recent historical year for which the operating model is conditioned on data (see **Table App.1.1**. for year definitions, years of MSE projection and years for which CMP advice is calculated)
Age class	There are 3 age classes that are used to model movement: Age class 1 consists of 1-4 year olds; age class 2 consists of 5-8 year olds; age class 3 consists of 9+ year olds.
Penalized maximum likelihood estimate (MLE)	The type of fitting (estimation) process used to condition operating models on data. The penalized MLE include priors (penalty functions) for parameters such as recruitment deviations in addition to the likelihood components for observed data. The M3 model was designed to operate as a Bayesian estimator using Markov Chain Monte Carlo and as such the penalized MLE is equivalent to the maximum posterior density (mode of the joint posterior distribution).

1.1 Spatial definitions

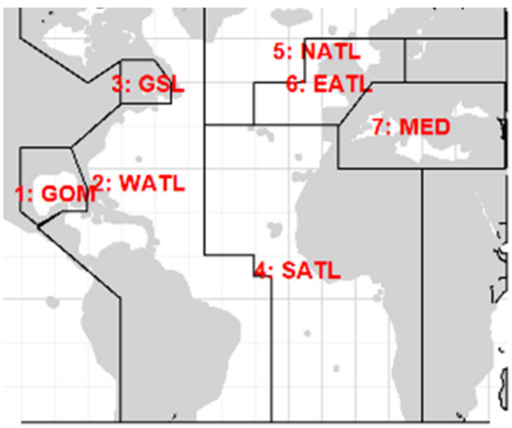

Figure 1.1. The seven spatial strata.

Baseline

The 7-stratum model of **Figure 1.1** (the reported electronic tagging data and the stock of origin data do not have sufficient resolution to divide the Mediterranean stratum).

Alternative low priority future options

The MAST model (Taylor et al. 2011) which has strata similar to **Figure 1.1**, but where strata 4-6 are merged into a single East Atlantic stratum.

1.2 Stock mixing

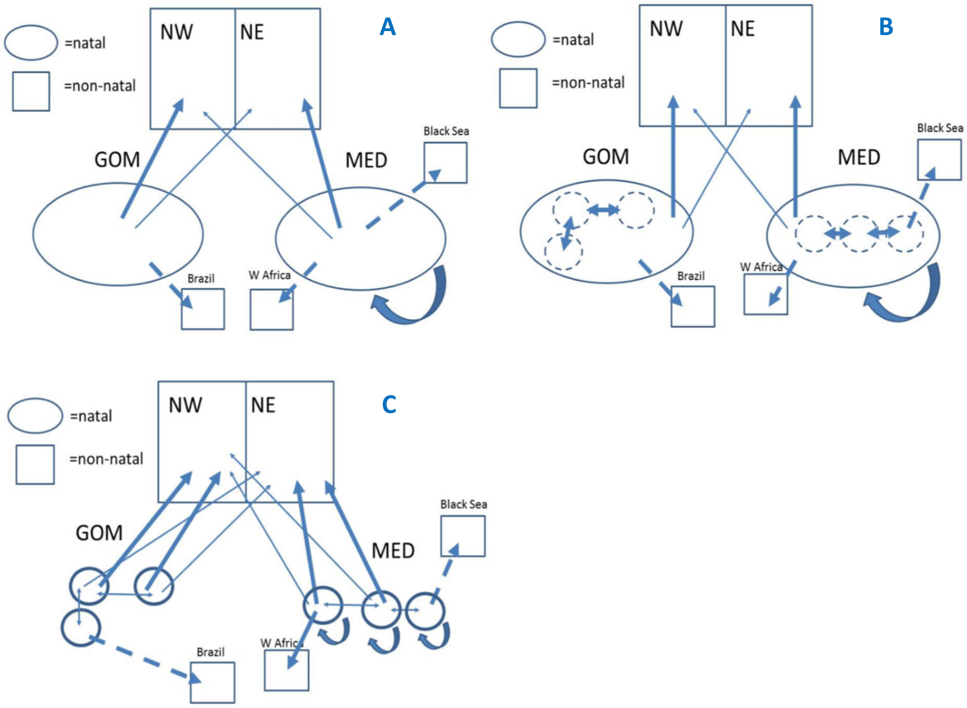

Figure 1.2. Mixing hypotheses suggested by Anon. (2014) and Arrizabalaga et al. (2019). (A) A two-stock model with no sub-stock structure. (B) A two-stock model with sub-stock structure. (C) A two-stock model with sub-stock structure.

<u>Baseline</u>

A two-stock model similar to **Figure 1.2-A** but adhering to the spatial structure of **Figure 1.1**. The mixing proportions are determined by the stock of origin data (genetics and otolith chemistry).

2 PAST DATA AVAILABLE

Table 2.5 provides an overview of the data that may be used to condition operating models for Atlantic bluefin tuna. The table indicates those data that have been gathered, those that are currently available and those that have already been used in conditioning operating models.

2.1 Raw data

Operating models are fitted to the fishery, tagging and survey data that are currently available (**Table 2.5**, field 'Used in OM'). Currently, the operating model is fitted to ICCAT Task II landings data scaled upwards to annual Task I landings.

The ICCAT catch-at-size dataset was used to estimate gear selectivity for each of the baseline fleet types. The operating models incorporate 18 fishing fleets, as described in **Table 3.1**.

The electronic tagging data from several sources (NOAA, DFO, WWF, AZTI, UNIMAR, IEO, UCA, FEDERCOOPESCA, COMBIOMA, GBYP, IFREMER, Stanford University) have been compiled by NOAA (M. Lauretta) and used to estimate movements among spatial strata. Tag tracks were provided by the seven spatial strata. These are converted to strata-quarter records by the following rule: for each tag, its strata position in a quarter is assigned as the strata in which the tag spent the most days during that quarter (**Figure 2.1**).

In the model developed for movement, quarterly transitions between strata depend on stock (Eastern or Western) and age class (ages 1-4, 5-8, and 9+). Only tags that have either corresponding weight or length data can be assigned an age class (by cohort slicing) and can be used by the model. Similarly, only those tags that have entered either the Gulf of Mexico or the Mediterranean can be assigned a stock of origin. All other tags are removed and not used in the conditioning of operating models. The exception are tags released by AZTI in the Bay of Biscay, which are assumed to correspond to Eastern Stock of origin. By November 2018 data from 1,307 tags were available for the model; however only 598 tag transitions (quarterly movements) could be used to provide transition information as the others lacked either age-class assignment or stock of origin assignment (**Figure 2.2**).

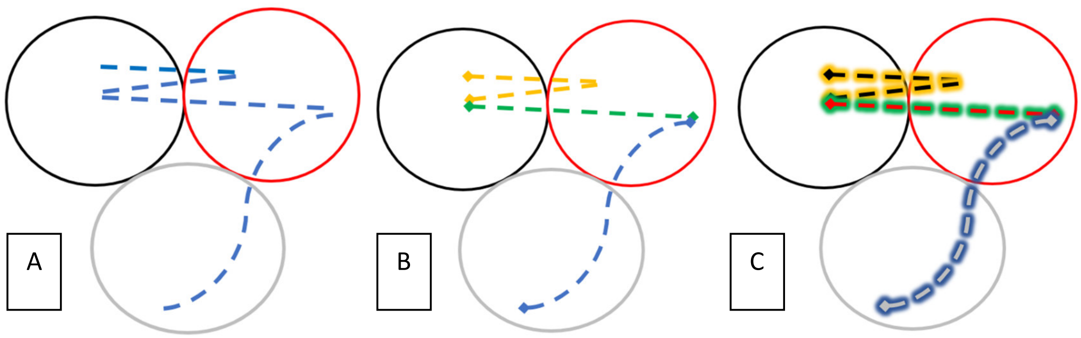

Figure 2.1. Electronic tag data were used to inform quarterly transitions. This figure explains how each tag was allocated to a stratum (represented as a black, red or grey circle) and quarter. The blue dashed line in (A) represents one electronic tag track. In (B) this track is sliced into quarters (here the track is split into different quarters through different colours 1=yellow, 2=green, 3=blue). Then (C) the track for each quarter was allocated to a spatio-temporal stratum (a spatial stratum, quarter, age class). This was done by counting the days (days are represented as dashes in these figures) the tag spent in each spatio-temporal stratum; the stratum where the tag spent the most days in a quarter was determined to be the location for the tag in that quarter.

(A) All age classes

TO AREA

FROM AREA

From Q4 / To Q1	Western Stock							Eastern Stock						
	GOM	WATL	GSL	SATL	NATL	EATL	MED	GOM	WATL	GSL	SATL	NATL	EATL	MED
GOM	7	0	0	0	0	0	0	0	0	0	0	0	0	0
WATL	44	45	0	0	0	0	0	0	12	0	1	0	0	0
GSL	1	1	0	0	0	0	0	0	0	0	0	0	0	0
SATL	0	0	0	0	0	0	0	0	0	0	4	0	4	1
NATL	0	0	0	0	0	0	0	0	2	0	3	1	0	0
EATL	0	0	0	0	0	0	0	0	1	0	8	0	9	0
MED	0	0	0	0	0	0	0	0	0	0	3	0	0	37

From Q1 / To Q2	GOM	WATL	GSL	SATL	NATL	EATL	MED	GOM	WATL	GSL	SATL	NATL	EATL	MED
GOM	37	4	0	0	0	0	0	0	0	0	0	0	0	0
WATL	28	12	0	0	0	0	0	0	8	0	6	2	1	2
GSL	0	0	0	0	0	0	0	0	0	0	0	0	0	0
SATL	0	0	0	0	0	0	0	0	1	0	8	0	4	1
NATL	0	0	0	0	0	0	0	0	0	0	1	0	0	0
EATL	0	0	0	0	0	0	0	0	0	0	2	0	5	1
MED	0	0	0	0	0	0	0	0	0	0	0	0	0	17

From Q2 / To Q3	GOM	WATL	GSL	SATL	NATL	EATL	MED	GOM	WATL	GSL	SATL	NATL	EATL	MED
GOM	2	18	3	0	0	0	0	0	0	0	0	0	0	0
WATL	0	18	2	0	0	0	0	0	7	0	1	0	1	0
GSL	0	0	0	0	0	0	0	0	0	0	0	0	0	0
SATL	0	0	0	0	0	0	0	0	1	0	2	2	3	3
NATL	0	0	0	0	0	0	0	0	0	0	0	1	1	0
EATL	0	0	0	0	0	0	0	0	0	0	1	0	5	0
MED	0	0	0	0	0	0	0	0	0	1	6	3	9	34

From Q3 / To Q4	GOM	WATL	GSL	SATL	NATL	EATL	MED	GOM	WATL	GSL	SATL	NATL	EATL	MED
GOM	0	0	0	0	0	0	0	0	0	0	0	0	0	0
WATL	1	18	2	0	0	0	0	0	7	0	0	0	0	0
GSL	6	29	0	0	0	0	0	0	2	0	0	0	0	0
SATL	0	0	0	0	0	0	0	0	0	0	4	0	0	1
NATL	0	0	0	0	0	0	0	0	0	0	0	4	0	0
EATL	0	0	0	0	0	0	0	0	0	0	1	4	22	0
MED	0	0	0	0	0	0	0	0	0	0	2	0	0	47

TOTAL 598

(B) Age class 1

TO AREA

FROM AREA

From Q4 / To Q1	Western Stock							Eastern Stock						
	GOM	WATL	GSL	SATL	NATL	EATL	MED	GOM	WATL	GSL	SATL	NATL	EATL	MED
GOM	0	0	0	0	0	0	0	0	0	0	0	0	0	0
WATL	0	0	0	0	0	0	0	0	1	0	0	0	0	0
GSL	0	0	0	0	0	0	0	0	0	0	0	0	0	0
SATL	0	0	0	0	0	0	0	0	0	0	3	0	0	0
NATL	0	0	0	0	0	0	0	0	0	0	0	0	0	0
EATL	0	0	0	0	0	0	0	0	1	0	5	0	8	0
MED	0	0	0	0	0	0	0	0	0	0	1	0	0	6

From Q1 / To Q2	GOM	WATL	GSL	SATL	NATL	EATL	MED	GOM	WATL	GSL	SATL	NATL	EATL	MED
GOM	0	0	0	0	0	0	0	0	0	0	0	0	0	0
WATL	0	0	0	0	0	0	0	0	1	0	0	1	0	0
GSL	0	0	0	0	0	0	0	0	0	0	0	0	0	0
SATL	0	0	0	0	0	0	0	0	1	0	4	0	2	1
NATL	0	0	0	0	0	0	0	0	0	0	0	0	0	0
EATL	0	0	0	0	0	0	0	0	0	0	1	0	3	0
MED	0	0	0	0	0	0	0	0	0	0	0	0	0	3

From Q2 / To Q3	GOM	WATL	GSL	SATL	NATL	EATL	MED	GOM	WATL	GSL	SATL	NATL	EATL	MED
GOM	0	0	0	0	0	0	0	0	0	0	0	0	0	0
WATL	0	0	0	0	0	0	0	0	2	0	0	0	0	0
GSL	0	0	0	0	0	0	0	0	0	0	0	0	0	0
SATL	0	0	0	0	0	0	0	0	0	0	0	0	0	0
NATL	0	0	0	0	0	0	0	0	0	0	0	0	1	0
EATL	0	0	0	0	0	0	0	0	0	0	0	0	2	0
MED	0	0	0	0	0	0	0	0	0	0	0	0	0	0

From Q3 / To Q4	GOM	WATL	GSL	SATL	NATL	EATL	MED	GOM	WATL	GSL	SATL	NATL	EATL	MED
GOM	0	0	0	0	0	0	0	0	0	0	0	0	0	0
WATL	0	0	0	0	0	0	0	0	1	0	0	0	0	0
GSL	0	0	0	0	0	0	0	0	0	0	0	0	0	0
SATL	0	0	0	0	0	0	0	0	0	0	1	0	0	1
NATL	0	0	0	0	0	0	0	0	0	0	0	0	0	0
EATL	0	0	0	0	0	0	0	0	0	0	0	0	18	0
MED	0	0	0	0	0	0	0	0	0	0	1	0	0	5

TOTAL 74

Figure 2.2. Observed electronic tag transitions among spatial strata by stock and quarter. These are tags present in a particular stratum (row) that move to a stratum (column) in the following quarter, and these transitions are derived from 1307 individual tags (A). The solid line represents strata available to the Western Stock (i.e., excludes the MED), the dashed line represents strata available to the Eastern Stock (i.e., excludes the GOM). The shaded diagonal cells highlight tags that did not move strata from one quarter to the next. (B) Age class 1, (C) age class 2, (D) age class 3 consist of 1-4 year olds, 5-8 year olds, 9+ year olds, respectively.

(C) Age class 2

TO AREA

FROM AREA

Western Stock — **From Q4** **To Q1**

	GOM	WATL	GSL	SATL	NATL	EATL	MED
GOM	0	0	0	0	0	0	0
WATL	0	1	0	0	0	0	0
GSL	0	0	0	0	0	0	0
SATL	0	0	0	0	0	0	0
NATL	0	0	0	0	0	0	0
EATL	0	0	0	0	0	0	0
MED	0	0	0	0	0	0	0

Eastern Stock — **From Q4** **To Q1**

	GOM	WATL	GSL	SATL	NATL	EATL	MED
GOM	0	0	0	0	0	0	0
WATL	0	0	0	0	0	0	0
GSL	0	0	0	0	0	0	0
SATL	0	0	0	1	0	0	1
NATL	0	0	0	0	0	0	0
EATL	0	0	0	0	0	0	0
MED	0	0	0	2	0	0	30

Western Stock — **From Q1** **To Q2**

	GOM	WATL	GSL	SATL	NATL	EATL	MED
GOM	0	0	0	0	0	0	0
WATL	0	2	0	0	0	0	0
GSL	0	0	0	0	0	0	0
SATL	0	0	0	0	0	0	0
NATL	0	0	0	0	0	0	0
EATL	0	0	0	0	0	0	0
MED	0	0	0	0	0	0	0

Eastern Stock — **From Q1** **To Q2**

	GOM	WATL	GSL	SATL	NATL	EATL	MED
GOM	0	0	0	0	0	0	0
WATL	0	0	0	0	0	0	0
GSL	0	0	0	0	0	0	0
SATL	0	0	0	1	0	0	0
NATL	0	0	0	0	0	0	0
EATL	0	0	0	0	0	0	0
MED	0	0	0	0	0	0	14

Western Stock — **From Q2** **To Q3**

	GOM	WATL	GSL	SATL	NATL	EATL	MED
GOM	0	0	0	0	0	0	0
WATL	0	2	0	0	0	0	0
GSL	0	0	0	0	0	0	0
SATL	0	0	0	0	0	0	0
NATL	0	0	0	0	0	0	0
EATL	0	0	0	0	0	0	0
MED	0	0	0	0	0	0	0

Eastern Stock — **From Q2** **To Q3**

	GOM	WATL	GSL	SATL	NATL	EATL	MED
GOM	0	0	0	0	0	0	0
WATL	0	0	0	0	0	0	0
GSL	0	0	0	0	0	0	0
SATL	0	0	0	1	0	0	0
NATL	0	0	0	0	0	0	0
EATL	0	0	0	0	0	0	0
MED	0	0	0	0	0	0	5

Western Stock — **From Q3** **To Q4**

	GOM	WATL	GSL	SATL	NATL	EATL	MED
GOM	0	0	0	0	0	0	0
WATL	0	1	0	0	0	0	0
GSL	0	0	0	0	0	0	0
SATL	0	0	0	0	0	0	0
NATL	0	0	0	0	0	0	0
EATL	0	0	0	0	0	0	0
MED	0	0	0	0	0	0	0

Eastern Stock — **From Q3** **To Q4**

	GOM	WATL	GSL	SATL	NATL	EATL	MED
GOM	0	0	0	0	0	0	0
WATL	0	0	0	0	0	0	0
GSL	0	0	0	0	0	0	0
SATL	0	0	0	0	0	0	0
NATL	0	0	0	0	0	0	0
EATL	0	0	0	0	0	0	0
MED	0	0	0	1	0	0	38

TOTAL **100**

(D) Age class 3

TO AREA

FROM AREA

Western Stock — **From Q4** **To Q1**

	GOM	WATL	GSL	SATL	NATL	EATL	MED
GOM	7	0	0	0	0	0	0
WATL	44	44	0	0	0	0	0
GSL	1	1	0	0	0	0	0
SATL	0	0	0	0	0	0	0
NATL	0	0	0	0	0	0	0
EATL	0	0	0	0	0	0	0
MED	0	0	0	0	0	0	0

Eastern Stock — **From Q4** **To Q1**

	GOM	WATL	GSL	SATL	NATL	EATL	MED
GOM	0	0	0	0	0	0	0
WATL	0	11	0	1	0	0	0
GSL	0	0	0	0	0	0	0
SATL	0	0	0	0	0	4	0
NATL	0	2	0	3	1	0	0
EATL	0	0	0	3	0	1	0
MED	0	0	0	0	0	0	1

Western Stock — **From Q1** **To Q2**

	GOM	WATL	GSL	SATL	NATL	EATL	MED
GOM	37	4	0	0	0	0	0
WATL	28	10	0	0	0	0	0
GSL	0	0	0	0	0	0	0
SATL	0	0	0	0	0	0	0
NATL	0	0	0	0	0	0	0
EATL	0	0	0	0	0	0	0
MED	0	0	0	0	0	0	0

Eastern Stock — **From Q1** **To Q2**

	GOM	WATL	GSL	SATL	NATL	EATL	MED
GOM	0	0	0	0	0	0	0
WATL	0	7	0	6	1	1	2
GSL	0	0	0	0	0	0	0
SATL	0	0	0	3	0	2	0
NATL	0	0	0	1	0	0	0
EATL	0	0	0	1	0	2	1
MED	0	0	0	0	0	0	0

Western Stock — **From Q2** **To Q3**

	GOM	WATL	GSL	SATL	NATL	EATL	MED
GOM	2	18	3	0	0	0	0
WATL	0	16	2	0	0	0	0
GSL	0	0	0	0	0	0	0
SATL	0	0	0	0	0	0	0
NATL	0	0	0	0	0	0	0
EATL	0	0	0	0	0	0	0
MED	0	0	0	0	0	0	0

Eastern Stock — **From Q2** **To Q3**

	GOM	WATL	GSL	SATL	NATL	EATL	MED
GOM	0	0	0	0	0	0	0
WATL	0	5	0	1	0	1	0
GSL	0	0	0	0	0	0	0
SATL	0	1	0	1	2	3	3
NATL	0	0	0	0	1	0	0
EATL	0	0	0	1	0	3	0
MED	0	0	1	6	3	9	29

Western Stock — **From Q3** **To Q4**

	GOM	WATL	GSL	SATL	NATL	EATL	MED
GOM	0	0	0	0	0	0	0
WATL	1	17	2	0	0	0	0
GSL	6	29	0	0	0	0	0
SATL	0	0	0	0	0	0	0
NATL	0	0	0	0	0	0	0
EATL	0	0	0	0	0	0	0
MED	0	0	0	0	0	0	0

Eastern Stock — **From Q3** **To Q4**

	GOM	WATL	GSL	SATL	NATL	EATL	MED
GOM	0	0	0	0	0	0	0
WATL	0	6	0	0	0	0	0
GSL	0	2	0	0	0	0	0
SATL	0	0	0	3	0	0	0
NATL	0	0	0	0	4	0	0
EATL	0	0	0	1	4	4	0
MED	0	0	0	0	0	0	4

TOTAL **424**

Figure 2.2. Continued.

Catch data provide scale to stock assessments. It follows that spatial stock of origin data are necessary to estimate the relative magnitude of the various stocks in a multi-stock model (to correctly assign catches to stock). Currently the model uses stock of origin data derived from the otolith microchemistry and genetic research of AZTI, UMCES, GBYP, and DFO (**Tables 2.5** to **2.9**).

There is uncertainty in regard to the stock of origin of bluefin tuna catches in the South Atlantic that were reported prior to 1970. For the Baseline, these are dealt with in the same way as all other catches: they are assigned to the strata of Figure 1.1 by uprating Task II catches (that are reported spatially) to the annual Task 1 catch data.

2.2 Analysed data

The operating models are also fitted to standardized CPUE indices (**Table 2.1**) and a range of fishery-independent indices (**Table 2.2**). These fishery-independent indices include a Western larval index in the Gulf of Mexico (GOM, Lauretta et al., 2018), an Eastern larval index in the western Mediterranean (Alvarez-Berastegui et al., 2020) and two aerial surveys in the Mediterranean (French Aerial survey: Rouyer et al., 2021, and GBYP aerial survey: Vázquez Bonales et al., 2018).

In order to predict observed catch at size from model predicted catch at age, operating models have made use of an inverse age-at-length key (probability of length class given age). These keys are developed from the base-case stock assessment growth curves for Eastern and Western stocks and a coefficient of variation (variability in length at age) determined by the growth model of Allioud et al. (2017).

Table 2.1. The standardized CPUE indices used to fit the operating models (many of which are used in stock assessments previously conducted by ICCAT). Many of these indices are available after 2019 but the operating mode uses data only to 2019 due to the unavailability of CATDIS updated catch data for more recent years at the time of model conditioning. The right-most column indicates the fishing fleets used to assign selectivity to each CPUE index; the fishing fleets are described in **Table 3.1**.

	Flag	Gear	Details	Fleet (selectivity) assigned
1	Spain	Baitboat	1952-2006, Q3, E Atl	3: BBold
2	Spain / France	Baitboat	2007-2014, Q3, E Atl	4: BBnew
3	Morocco / Spain	Trap	1981-2011, Q2, S Atl	12: TPold
4	Morocco / Portugal	Trap	2012-2020, Q2, S Atl	13: TPnew
5	Japan	Longline	1975-2009, Q2, S Atl	2: LLJPN
6	Japan	Longline	1990-2009, Q4, NE Atl	2: LLJPN
7	Japan	Longline	2010-2019, Q4, NE Atl	18: LLJPNnew
8*	US (66cm - 114cm)	Rod and reel	1995-2020, Q3, W Atl	15: RRUSAFS (50 –125cm)
9*	US (115cm - 144cm)	Rod and reel	1995-2020, Q3, W Atl	15: RRUSAFS (100 – 150cm)
10	US (177cm+)	Rod and reel	1993-2020, Q3, W Atl	16: RRUSAFB (175cm+)
11	US (<145cm)	Rod and reel	1980-1992 (gap in 1984), Q3, W Atl	15: RRUSAFS (50 – 150cm)
12	US (195cm+)	Rod and reel	1983-1992, Q3, W Atl	16: RRUSAFB (200cm+)
13	Japan	Longline	1974-1980, Q2, GOM	2: LLJPN
14	Japan	Longline	1976-2009, Q4, W Atl	2: LLJPN
15	Japan	Longline	2010-2020, Q4, W Atl	18: LLJPNnew
16	Canada GSL	Rod and reel	1988-2020, Q3, GSL	14: RRCAN
17	Canada SWNS	Rod and reel	1996-2020, Q3, W Atl	14: RRCAN
18	US (66cm – 144cm)	Rod and reel	1995-2020, Q3, W Atl	15: RRUSAFS (50 – 150cm)
19	US-Mexico	Longline	1994-2019, Q2, GOM	1: LLOTH

** US RR indices for 66cm – 114cm and 115cm – 144cm would be replaced by US RR index for 66cm – 144cm (index 18)*

Table 2.2. Fishery-independent indices used in the fitting of operating models.

	Type	Details	Infers:
1	Canadian acoustic survey past	1994-2017, Q3, GSL, index in number of fish	Number of combined eastern and western fish in Q3 for the GSL stratum according to the estimated vulnerable biomass available to the CANRR fleet for 150cm plus
2	Canadian acoustic survey recent **(zero weight)**	2017-2018, Q3, GSL, index in number of fish	
3	French aerial survey past	2000-2003, Q3, Med	Vulnerable biomass in Q3 in Med, according to the RRUSAFS selectivity which is used due to similar assumed size of fish
4	French aerial survey recent	2009-2019 (gap in 2013), Q3, Med	
5	Aerial survey – GBYP* **(zero weight)**	2010-2018 (gaps in 2012, 2014, and 2016), Q2, Med	SSB eastern stock in Q2 in Med
6	USA GOM Larval survey	1977-2019 (gaps in 1979-1980, and 1985), Q2, GOM	SSB western stock in Q2 in GOM stratum
7	Western Med Larval survey	2001-2019 (gaps in 2006, 2007, 2009, 2011), Q2, Med	SSB eastern stock in Q2 in Med

** Only the Balearic component is used for SSB (because there are problems with consistency regarding patchy or low biomass inference in other strata surveyed). Note that this survey awaits update, hence the zero weight at this time.*

In order to initialize the spatial-seasonal operating model at a plausible distribution of vulnerable biomass, a so-called "master index" was derived. This index allows a standardized effort to be derived for the catch series of any fleet, thereby simplifying the estimation of fishing mortality rates (for more detail see Carruthers and Butterworth 2018). The only role of the master index is to initialize the model and it effectively plays no role in the likelihood.

The default master index ('Assess-Tag') was derived using electronic tagging data and East / West area trends estimated by the 2017 Stock Synthesis assessments (Anon. 2017).

The electronic tagging data for known stock of origin s (fish that have been in either the Gulf of Mexico or the Mediterranean) were aggregated by quarter m, stratum-from a, stratum-to k, into a matrix T. Each row of this matrix was normalized to form a Markov movement matrix V such that the values summed to 1:

$$V_{s,m,a,k} = \frac{T_{s,m,a,k}}{\sum_k T_{s,m,a,k}} \qquad (2.1)$$

For each stock, an even initial spatial distribution was repeatedly multiplied though this quarterly movement matrix until it stabilized on an asymptotic quarterly distribution w (Eq. 3.14).

Then using the estimates of historical spawning stock biomass B from the 2017 East and West Stock Synthesis stock assessments (assuming that the area trends of assessments reflect the stock trends), a predicted spawning biomass by season and stratum $\hat{B}$ was calculated:

$$\hat{B}_{s,m,a,y} = w_{s,m,a} B_{s,y} \qquad (2.2)$$

This was summed over the two stocks (Eastern and Western) to get total biomass $\bar{B}$ that was assumed to be proportional to the master index I (red line **Figure 2.3**):

$$I_{m,a,y}^{Assess-Tag} \approx \bar{B}_{m,a,y} = \sum_s \hat{B}_{s,m,a,y} \qquad (2.3)$$

Other approaches to index derivation have been used to demonstrate that it has little impact on final model estimates (Carruthers and Butterworth, 2018) including derivation by GLM and assuming a flat, constant trend over time and space (**Figure 2.3**). Note that in order to update the Master index to include years after 2015 (the last year that VPA and Stock Synthesis assessment SSB estimates have been available for both stocks), the final assessment SSB value of 2015 was assumed for all future years.

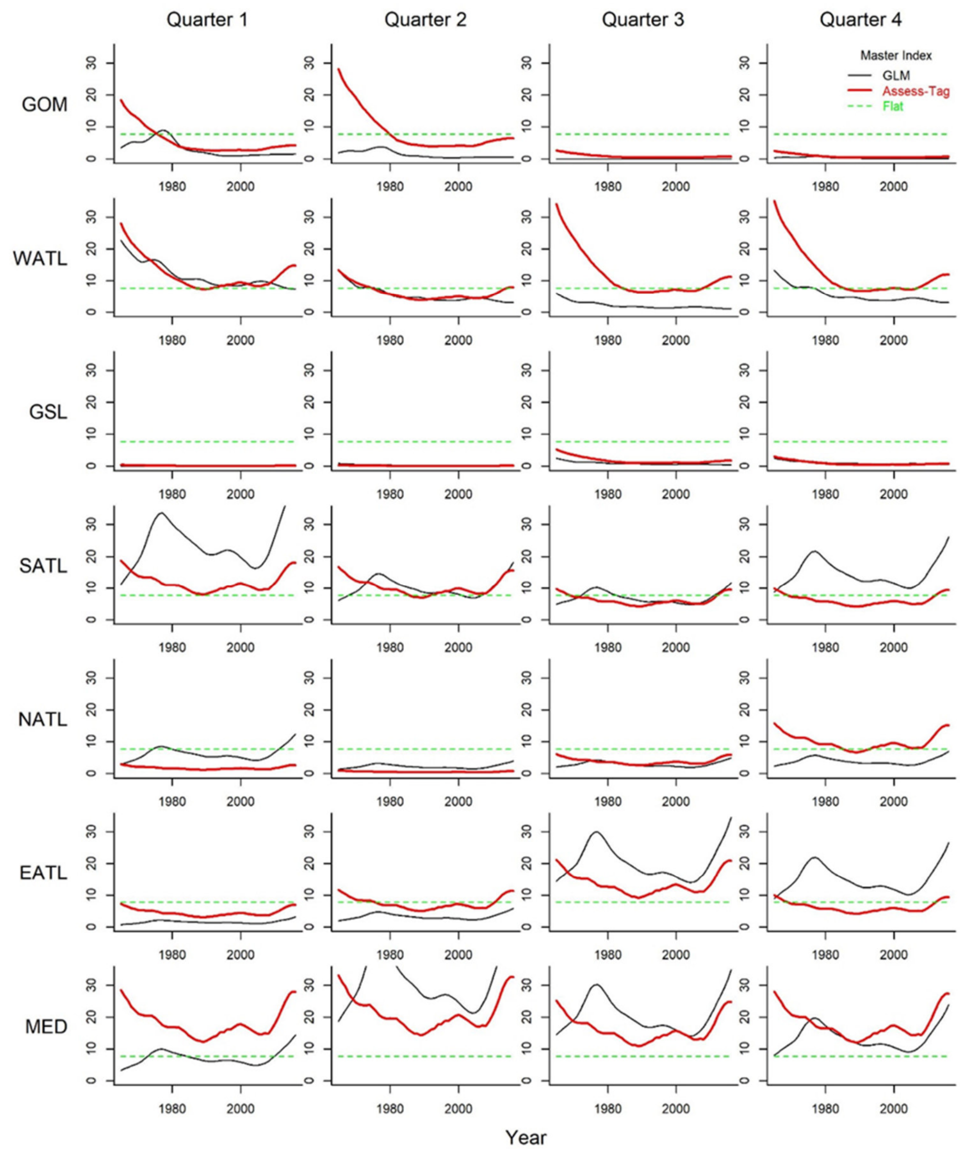

Figure 2.3. The seasonal / spatial master indices derived by various methods including by generalized linear modelling (GLM, Carruthers 2018) and 'Assess-Tag' described above and in Carruthers and Butterworth (2018). The master index is used as a way to initialize the model and has essentially no other role in model fitting. The default index used in OM conditioning is Assess-Tag described above but results have been shown to be invariant to the choice of master index. The magnitude of all points are relative and have an arbitrary mean value that is shown by the green dashed line.

2.3 Assumptions

The following are the default assumptions made in all operating models:

- The age-length key is static and not adjusted according to fishing mortality rate and length selectivity of fishing.

- Larval indices are assumed to be proportional to spawning stock biomass in the stratum in which they were collected, in contrast to stock-wide spawning stock biomass (for scenarios where the two are not proportional). Non-larval fishery-independent indices may mirror fleet selectivities or are assumed to be proportional to spawning stock biomass.

- Fish found in the GOM stratum are assumed to be all Western Stock fish (i.e., the model assumes that effectively no Eastern Stock fish move to the GOM). Fish found the MED stratum are all assumed to be Eastern Stock fish.

With the exception of one robustness set OM specification all CPUE indices are considered to be proportional to vulnerable (i.e., selectivity-weighted) biomass.

Table 2.5. Overview of data that may (includes all available years, not just those used in conditioning) be used to inform operating models for Atlantic bluefin tuna (available online here). Cells shaded green reflect sources for which data are available ('Collab', the Technical Team TT, or the ICCAT secretariat) and whether data that are available have also been used in conditioning preliminary operating models ('used in OM?'). Conventional tags are used only in defining the stock-specific areas of the GOM and Mediterranean.

Type of data (Informs)	Year range	Til	Spatial range	Can be by quarter?	By age-class?	Contact	Collab	Available to: TC	TT	ICCAT	ALL	Used in OM?
1. CPUE indices (relative abundance, movement, performance at stakeholder level)												
1.1. Japanese LL standardized spatial	1990-2019	∞	NE	Y	N	Y. Tsukahara	Y	Y	Y	Y	Y	Y
	1975-2019	∞	W	Y	N		Y	Y	Y	Y	Y	Y
	1976-2010	2010	E, M	Y	N		Y	Y	Y	Y	Y	Y
	1974-1980	1980	GOM	Y	N		Y	Y	Y	Y	Y	Y
1.2. USA-MEX LL standardized spatial	1994-2019	∞	GOM	Y	N	M. Lauretta	Y	Y	Y	Y	Y	Y
1.3. USA RR standardized spatial	1993-2019	∞	W	Y	N		Y	Y	Y	Y	Y	Y
1.4. CAN RR GSL standardized	1984-2019	∞	GSL	Y	N	A. Hanke	Y	Y	Y	Y	Y	Y
1.5. CAN RR SWNS standardized	1988-2019	∞	W	Y	N		Y	Y	Y	Y	Y	Y
1.6. MOR-SPN TRAP standardized	1982-2011	2011	WM	Y	N	N. Abid, J. Ortiz de	Y	Y	Y	Y	Y	Y
1.7. MOR-POR TRAP standardized	2012-2020	∞	E, WM	Y	N	N. Abid, P. Lino	Y	Y	Y	Y	Y	Y
1.8. SPN BB	1952-2006	2006	EATL	Y	N	H. Arrizabalaga	Y	Y	Y	Y	Y	Y
1.9. SPN-FR BB	2007-2014	2014	EATL	Y	N	H. Arrizabalaga	Y	Y	Y	Y	Y	Y
2. Fish. Ind. surveys (relative abundance, movement)												
2.1. GBYP Aerial	2010-2018	∞	M	Y	N	GBYP	Y	Y	Y	Y	Y	Y
2.2. French Aerial	'00-'03, '09-'19	∞	M	Y	N	T. Rouyer	Y	Y	Y	Y	Y	Y
2.3. Can Acoustic	'94-'17, '17-'18	∞	W	Y	N	T. Minch	Y	Y	Y	Y	Y	Y
3. Larval indices (SSB, movement)												
3.1. USA	1977-2019	∞	GOM	Y	N	W. Ingram	Y	Y	Y	Y	Y	Y
3.2. Western Med	'01-'05 '12-'19	2018	W Med	Y	N	GBYP	Y	Y	Y	Y	Y	Y
4. Catches (stock size, harvest rate)												
4.1. ICCAT task I		∞	non-spatial	N	N		Y	Y	Y	Y	Y	N
4.2. ICCAT task II	1950-2019		All	Y	N	ICCAT STAT	Y	Y	Y	Y	Y	N
4.3. ICCAT CATDIS					N		Y	Y	Y	Y	Y	Y
4.4 GBYP	1512-1950		E, M	Y	N		Y	Y	Y	Y	Y	Y
5. Catch composition (selectivity, depletion)												
5.1. ICCAT catch-at-size	1950-2019	∞	All	Y	N		Y	Y	Y	Y	Y	Y
5.2. Stereo video caging	2014	ended	WM, EM	Y	N	ICCAT STAT	N	N	N	N	N	N
5.3. GBYP Historical catches	1910-1950	=	E, M	Y	N		Y	Y	Y	Y	Y	Y
6. Conventional tags (feasible movement, growth, GTG heterogeneity)												
6.1. ICCAT	1954-2014	2015	All	Y	Y	ICCAT STAT	Y	Y	Y	Y	Y	Stock defs
7. Electronic tags (movement)												
7.1. LPRC (n=316)	2005-2009	ended	W	Y	Y	M. Lutcavage	Y	Y	N	N	N	Y
7.2. DFO (n=89)	2009-2018	∞	GSL,W,GOM	Y	Y	A. Hanke	Y	Y	N	N	N	Y
7.3. Stanford (n=391)	1996-2015	∞	All	Y	Y	B. Block	Y	Y	N	N	N	Y
7.4. GBYP (n=176)	2011-2017	2015	E, M	Y	Y	GBYP	Y	Y	N	N	N	Y
7.5. WWF (n=86)	2008-2015	2015	All	Y	Y	P. Cermeno	Y	Y	N	N	N	Y
7.6. NOAA (n=31)	2010-2013	2013	GOM,W,GSL	Y	Y	C. Brown	Y	Y	N	N	N	Y
7.7. DFO-Acadia (n=37)	2010-2011	ended	GSL	Y	Y	M. Stokesbury	Y	Y	N	N	N	Y
7.8. UCA (n=46)	2009-2011	ended	W, SE, NE, E, I	Y	Y	A. Medina	Y	Y	N	N	N	Y
7.9. DFO - Duke (n = 15)	2007-2008	ended	W	Y	Y	A. Hanke	Y	Y	N	N	N	Y
7.10. IEO (n=13)	2001	ended	SE, E, NE, M	Y	Y	F. Abascal	Y	Y	N	N	N	Y
7.11. IFREMER (n=47)			SE, E, M	Y	Y	T. Rouyer	Y	Y	N	N	N	Y
7.12. AZTI (n=20)	2005-2011	ended	SE, E, NE, W	Y	Y	H. Arrizabalaga	Y	Y	N	N	N	Y
7.13. GBYP-Unimar (n=40)	2007-2015	ended	SE, E, NE, M	Y	Y	GBYP	Y	Y	N	N	N	Y
8. Otolith microchemistry (stock of origin)												
8.1. AZTI (n=189)	2009-2011	2011	E	Y	Y	I. Fraile	Y	Y	Y	Y	Y	Y
8.2. US+Can (n=3545)	1974-2015	∞	W, GSL	Y	Y	A. Hanke	Y	Y	Y	Y	Y	Y
8.3. GBYP (n=2237)	2009-2016	∞	All	Y	Y	GBYP	Y	Y	Y	Y	Y	Y
9. Otolith shape analysis (stock of origin)												
9.1. GBYP (n=172)	2011-2013	2015	E, W, WM	Y	N	GBYP	Y	N	N	N	N	N
10. SNP (population structure, genetic structure)												
10.1. Med HCMR					N	G. Zmpicinini	N	N	N	N	N	N
10.2. GBYP (n=789)	2011-2015	∞	All		N	GBYP	Y	Y	N	N	N	Y
10.3 NOAA/VIMS/CSIRO	2015	∞	GOM/M	N	N	J. Walter	N	N	N	N	N	N
10.4 GBYP Historical UB	200 BC - 1927	1927	E, M	Y	N	A. Cariani	Y	N	N	N	N	N

Table 2.5 Continued.

Type of data (Informs)	Year range	Til	Spatial range	Can be by quarter?	By age-class?	Contact	Collab	Available to: TC	TT	ICCAT	ALL	Used in OM?
11. Other genetics on population structure (population structure, genetic structure)												
11.1. mtDNA					N	B. Block	N	N	N	N	N	N
11.2. Micro Sat/ mtDNA (n=320 / 147)	2003	ended	GOM, WM	Y	N	Carlsson	N	N	N	N	N	N
12. Growth, aging (age-length keys, length-age keys)												
12.1. Age-length keys (NOAA)				Y	N	J. Walter	Y	N	N	N	N	N
12.2. Age-length keys (IEO)	2010-2012	ended	E, WM	Y	N	E. Rodriguez-Marin	Y	N	N	N	N	N
12.3. Age-length keys (DFO)	2010-2013	ended	GSL, W	Y	N	A. Hanke	Y	N	N	N	N	N
12.4. Derived from tagging	1963-2012	ended	Es, W s	Y	N	L. Allioud	Y	Y	N	N	N	Y
12.5 Age-length keys (GBYP)	2011-2015		E, M	Y	N	GBYP	Y	N	Y	Y	Y	N
12.6 Ageing calibration (GBYP)	2014		E, M	Y	N	GBYP	Y	N	Y	Y	Y	N
13. Maturity (Spawning biomass)												
13.1. Western (NOAA)	1975-1981	ended	GOM	Y	N	G. Diaz	Y	N	N	N	N	N
13.2 Mediterranean			M	Y	N	GBYP	Y	N	N	N	N	N
14. Other ecological data (spatial distribution, covariates for CPUE standardization, steepness, natural mortality rate, spawning locations etc.)												
14.1. Larval ecology (IEO)		ended	WM	Y	N	D. Alvarez Berastegui		N	N	N	N	N
14.2. Habitat model				Y	N	J-N. Druon		N	N	N	N	N

Table 2.6. Summary of the observed assignment scores from otolith microchemistry and genetics datasets (labelled 'Probability Eastern Origin' from dataset 'Joint East West Mixing Data 15042019.csv'). Each point in those datasets consists of an observed assignment score, i.e., the assigned probability (between 0% and 100%) of that point being of Eastern origin. The table summarises (median, 5th and 95th percentiles) the observed assignment scores in each spatial stratum. A mixture model is applied to these data (Carruthers and Butterworth 2018) to generate stock-of-origin "pseudo-observations" that are used in the conditioning of the operating models.

Type	Percentile	GOM	WATL	GSL	SATL	NATL	EATL	MED
Otolith microchemistry	5th	0%	1%	4%	14%	6%	48%	32%
	Median	**7%**	**27%**	**23%**	**87%**	**75%**	**87%**	**84%**
	95th	48%	97%	89%	99%	97%	98%	97%
Genetics	5th	0%	0%	0%	4%	9%	29%	40%
	Median	**0%**	**45%**	**56%**	**82%**	**96%**	**98%**	**99%**
	95th	94%	100%	100%	100%	100%	100%	100%

Table 2.7. The sample size of stock of origin data by type (otolith micro-chemistry and genetics) and the 7 spatial strata. Note that data are available for the Gulf of Mexico and the Mediterranean, but that these were not used directly in the operating model. Instead, they were used to identify a western and eastern stock signature for interpreting the assignment data in a mixture model (Carruthers and Butterworth 2018).

	WATL	GSL	SATL	NATL	EATL	Total	%
Otolith Chemistry	2518	864	257	315	251	4205	76.7%
Genetics	165	64	491	429	127	1276	23.3%

Table 2.8. Seasonal-spatial coverage of the otolith chemistry assignment data (that have covariate information regarding age class and quarter; from dataset 'Joint East West Mixing Data 15042019.csv'). Orange shaded cells represent quarter-strata for which there are no stock of origin data available for the mixture model approach (i.e., no otolith chemistry data were available for these spatio-temporal strata).

Quarter	WATL	GSL	SATL	NATL	EATL	Total	%
1: Jan-Mar	369	0	39	0	0	408	9.7%
2: Apr-Jun	310	0	155	0	14	479	11.4%
3: Jul-Sept	1534	604	33	4	216	2391	56.9%
4: Oct-Dec	305	260	30	311	21	927	22.0%
Total	2518	864	257	315	251		
%	59.9%	20.5%	6.1%	7.5%	6.0%		

Table 2.9. Seasonal-spatial coverage of the genetics assignment data (that have covariate information regarding age class and season; from dataset 'Joint East West Mixing Data 15042019.csv'). Orange shaded cells represent quarter-strata for which there are no stock of origin data available for the mixture model approach (i.e., no genetic data were available for these spatio-temporal strata). Note that data are available for the Gulf of Mexico and the Mediterranean, but that these were not used in the operating model.

Quarter	WATL	GSL	SATL	NATL	EATL	Total	%
1: Jan-Mar	0	0	105	0	0	105	8.2%
2: Apr-Jun	0	0	268	1	8	277	21.7%
3: Jul-Sept	109	43	53	193	118	516	40.4%
4: Oct-Dec	56	21	65	235	1	378	29.6%
Total	165	64	491	429	127		
%	12.9%	5.0%	38.5%	33.6%	10.0%		

3 BASIC OPERATING MODEL DYNAMICS

3.1 Overview

The current operating model (modifiable multi-stock model, 'M3' v7.0) is based on conventional age-structured accounting (e.g., Quinn and Deriso 1999, Chapter 8) which is common to stock assessment models such as Stock Synthesis 3 (Methot and Wetzel 2013), CASAL (Bull et al. 2012), Multifan-CL (Fournier et al. 1998) and iSCAM (Martell 2015).

The standard age-structured equations are complicated somewhat by the quarterly temporal structure, in which age increments and recruitment occur in a particular quarter. In this version of the model, spawning occurs for all stocks in quarter 2 (spawning in the Mediterranean for the Eastern stock and Gulf of Mexico and West Atlantic strata for the Western stock is assumed to occur after a period of movement early in the year).

3.2 Equations

Numbers of individuals *N*, for stock *s*, in a model year *y*, in the first quarter *m*=1, age class *a*, and stratum *r* are calculated from individuals that have moved $\vec{N}$ (defined Eq. 3.18), in the previous year, final quarter (m=4) of the same age class, subject to combined natural and fishing mortality rate *Z*:

$$N_{s,y,m=1,a,r} = \vec{N}_{s,y-1,m=4,a,r} \cdot e^{-Z_{s,y-1,m=4,a,r}} \tag{3.1}$$

where total mortality rate is calculated from annual natural mortality rate *M*, divided by the fraction of the year represented by the quarter t_m (i.e., 0.25), and fishing mortality rate *F* (per quarter), summed over all fleets *f*:

$$Z_{s,y,m,a,r} = t_m\, M_{s,a} + \sum_f F_{s,y,m,a,r,f} \tag{3.2}$$

Fishing mortality rate at age is derived from fishing mortality rate by length class[1] F_l and the conditional probability of fish being in length class *l*, given age *a* (an inverse age-length key, LAK):

$$F_{s,y,m,a,r,f} = \sum_l F_{s,y,m,l,r,f} \cdot LAK_{s,a,l} \tag{3.3}$$

The LAK was calculated from stock-specific growth curves and assumes that uncertainty (CV) in length is a linear function of length: $CV_L = a_{cv} \cdot L + b_{cv}$ (Allioud et al. 2017) (**Table 8.2** for parameter values).

[1] 25cm length bins were assumed in OM conditioning since the principal model estimates (e.g., spawning stock depletion, BMSY) were consistently invariant to finer disaggregation but also required a much larger computational overhead in the calculation of equation 3.3. Additionally, the factor of the reference grid relating to length data weighting spans a much wider range of uncertainty than the resolution of the length composition data.

The fishing mortality rate at length is calculated from an index of fishing mortality rate I (calculated from dividing the value of the catch for that fleet by the value of the 'master index' in that year-quarter-stratum – a simple way to preserve scale), an estimated catchability coefficient q, a quarter and stratum specific deviation F_D (constrained to mean 1), a quarter, stratum and year specific deviation F_A (constrained to mean 1), and a length selectivity function s, by fleet:

$$F_{l,y,m,l,r,f} = q_f \cdot I_{y,m,r,f} \cdot F_{D,m,r} \cdot F_{A,y,m,r} \cdot s_{f,l} \tag{3.4}$$

For most fleets, selectivity is calculated by a double-normal equation using the mean length L_l for a length class l:

$$s_{f,l} = \begin{cases} 2^{-\left(\frac{L_l - l_{max,f}}{\sigma_{f,A}{}^2}\right)^2} & L_l \le l_{max,f} \\ 2^{-\left(\frac{L_l - l_{max,f}}{\sigma_{f,D}{}^2}\right)^2} & L_l > l_{max,f} \end{cases} \tag{3.5}$$

where $l_{max,f}$ is the fleet-specific length at maximum vulnerability, and $\sigma_{f,A}$ and $\sigma_{f,D}$ are parameters controlling the width of the ascending and descending limbs of the selectivity, respectively. Large values of $\sigma_{f,D}$ approximate a 'flat topped' logistic selectivity.

To ensure numerical stability and prevent the estimation of unrealistic values, the length at maximum vulnerability l_{max} and the two standard deviation parameters, σ_A and σ_D were derived from estimated parameters θ_{lmax}, θ_A and θ_D, respectively, and the longest length class L_{nl} (mean length of the maximum length class), as defined below (all these terms are by fleet, but these equations drop the fleet subscript for simplicity):

$$l_{max} = \rho_L + (\rho_U - \rho_L) \cdot \left(\frac{1}{20} + \frac{19}{20} \cdot \frac{e^{\theta_{lmax}}}{1+e^{\theta_{lmax}}}\right) \tag{3.6}$$

$$\sigma_A = 2\, l_{max} \cdot \left(\frac{e^{\theta_A}}{1+e^{\theta_A}}\right) \tag{3.7}$$

$$\sigma_D = (\rho_U - \rho_L) \cdot e^{\theta_D} \tag{3.8}$$

The ρ terms are the upper (ρ_U) and lower bounds (ρ_L) (lengths) for the truncation of the length selectivity function.

These parameterizations allow for unbounded estimation of the θ parameters. Each of these parameters has an extremely weak "prior" (penalty function) prescribed which still allows the model to converge in extreme cases where there is little or no data to inform a parameter. For example, if data suggest there is asymptotic (near logistic) selectivity, l_{max} tends to L_{nl} and there are no data above this length class to estimate the descending limb parameter θ_D.

In general, age or length structured models are much better informed by the data if at least one fleet selectivity either has the descending limb parameters fixed or can be assumed to have the form of a logistic ('flat topped') ogive. Without such a constraint, the declining frequency of older/longer classes can be attributed to either mortality rates or dome-shaped selectivity, and this parameter confounding can lead to poorly defined estimation and numerical instability during fitting. In this case at least one fleet is assumed to have a 2-parameter logistic form for its selectivity function:

$$s_{f,l} = \frac{1}{1+e^{\left(l_{inf,f} - L_l\right)/\sigma_{S,f}}} \tag{3.9}$$

where l_{inf} is the inflection point (the length at 50% vulnerability) and σ_S is a slope parameter controlling how steeply selectivity increases with length. Similar to the 3-parameter double-normal function, there is a reparameterization to ensure numerical stability during fitting:

$$l_{inf} = \rho_L + (\rho_U - \rho_L) \cdot \left(\frac{1}{20} + \frac{17}{20} \cdot \frac{e^{\theta_{linf}}}{1+e^{\theta_{linf}}}\right) \tag{3.10}$$

$$\sigma_S = (\rho_U - \rho_L) \cdot \left(0.005 + 0.11 \cdot \frac{e^{\theta_S}}{1+e^{\theta_S}}\right) \tag{3.11}$$

Again, the estimation of the θ parameter values can be unbounded.

All selectivity θ parameters are assigned a vague normal "prior" with mean 0.

In the spawning quarter *ms*, and spawning strata *rs*, ages advance by one and recruitment occurs.
The model includes a plus group which is the final age class n_a:

$$N_{s,y,ms,a,r} = \begin{cases} \vec{N}_{s,y,ms-1,a-1,r} \cdot e^{-Z_{s,y,ms-1,a-1,r}} & 1 < a < n_a \\ \vec{N}_{s,y,ms-1,a-1,r} \cdot e^{-Z_{s,y,ms-1,a-1,r}} + \vec{N}_{s,y,ms,a,r} \cdot e^{-Z_{s,y,ms,a,r}} & a = n_a \end{cases} \tag{3.12}$$

Recruitment is calculated based on stock-wide spawning stock biomass, and recruits enter the model in proportion to spawning stock biomass in the spawning strata (Gulf of Mexico and West Atlantic for the Western stock, the Mediterranean for the Eastern stock) in the spawning season (quarter 2 for both stocks). The model does not force all the SSB back to the spawning stratum at the spawning time. It is not possible to estimate recruitment for only those fish that exist in the spawning stratum in the spawning season. as that leads to numerical instability and can result in unrealistic estimates of recruits per spawner.

Recruitment (fish in their first year) is calculated from a Beverton-Holt stock recruitment relationship with fixed steepness:

$$N_{s,y,ms,1,rs} = \exp\left(\varepsilon_{R,s,y} - \sigma_{R,s}^2/2\right) \cdot \frac{\frac{4}{5} \cdot h_s \cdot R_{0,s} \cdot SSB_{s,y}}{\frac{1}{5} \cdot SpR_s \cdot R_{0,s} \cdot (1-h_s) + (h_s - 0.2) \cdot SSB_{s,y}} \cdot \nabla_{rs} \tag{3.13}$$

where ε_R is a random normal deviate with variance σ_R^2 and $\sigma_R^2/2$ is the bias correction to ensure that on average over years, recruitment strengths have a mean of 1. The ∇_{rs} term is the model estimated fraction of spawning stock biomass distributed among designated spawning areas. In the East, there is only one spawning area, the Mediterranean ($\nabla_{rs} = 1$), but in the West spawning can occur in either the Gulf of Mexico or the West Atlantic and recruits are distributed by ∇_{rs}.

Spawning stock biomass *SSB*, is calculated from moved stock numbers in the previous year and quarter prior to the spawning quarter *ms*, weight of individuals at age *w*, and the fraction of individuals mature at age *m*:

$$SSB_{s,y} = \sum_a \sum_r \vec{N}_{s,y,ms-1,a,r} \cdot e^{-Z_{s,y,ms-1,a,r}} \cdot w_{s,a} \cdot m_{s,a} \tag{3.14}$$

where weight is calculated from length at age *l*:

$$w_{s,a} = \alpha_s \cdot l_{s,a}^{\beta_s} \tag{3.15}$$

and the fraction mature at age is assumed to be a logistic function of age with parameters for the age at 50% maturity γ, and slope ϑ:

$$m_{s,a} = 1/\left(1 + e^{(\gamma_s - a)/\vartheta_s}\right) \tag{3.16}$$

Stock numbers for quarters that are not the first quarter of the year and are not the spawning quarter are calculated as:

$$N_{s,y,m,a,r} = \vec{N}_{s,y,m-1,a,r} \cdot e^{-Z_{s,y,m-1,a,r}} \tag{3.17}$$

In each quarter, before mortality and recruitment, fish are moved according to an age-class-specific Markov transition matrix *mov* that represents the probability of a fish moving from stratum *k* to stratum *r* at the end of the quarter *m*:

$$\vec{N}_{s,y,m,a,r} = \sum_k N_{s,y,m,a,k} \cdot mov_{s,m,a,k,r} \tag{3.18}$$

The movement matrix is calculated from a log-space matrix *lnmov* and a logit model to ensure each row (k) sums to 1:

$$mov_{s,m,a,k,r} = e^{lnmov_{s,m,a,k,r}} / \sum_r e^{lnmov_{s,m,a,k,r}} \tag{3.19}$$

Size/age stratification for movement models has been implemented for three age groups: 1-4, 5-8 and 9+ years (this will be kept the same for the Western Atlantic and the Eastern Atlantic/Mediterranean, but should be re-evaluated for the East as future data become available).

Due to the relatively incomplete coverage (over stocks, quarters and spatial strata) of electronic tagging data to be able to explicitly inform individual movements to/from each stratum, a parsimonious gravity modelling approach was used to estimate movement (e.g., MAST Taylor et al., 2011, Carruthers et al., 2011). For a movement matrix of dimension n_{strata} x n_{strata}, rather than estimating a parameter for each possible transition (which would result in (n_{strata} -1) x n_{strata} parameters), the gravity model estimates only the attractivity g of each stratum (n_{strata}-1 parameters) identically for all strata of departure. Unmodified, this is simply a spatial distribution model, mixing all tagged fish in every time step and redistributing them by fractions over all strata. There is however evidence of stock viscosity, where fish remain in the same strata over several time steps. This is particularly the case for spawning strata in spawning seasons, for example. The gravity model incorporates a single additional parameter per movement matrix (resulting in n_{strata} parameters per movement matrix) that is added to the positive diagonal (probability of staying in the same stratum, i.e., when the 'from strata' k is the same as the 'to strata' r) to make fish more likely to stay in proportion to the attractivity of that stratum:

$$lnmov_{s,m,2,k,r} = \begin{cases} g_{s,m,2,r} & k \neq r \\ g_{s,m,2,r} + e^{v_{m,2}} & k = r \end{cases} \tag{3.20}$$

In this equation, the subscript "2" refers to the second movement age-class (ages 5-8). Since the *lnmov* variables are used in a logit model to determine fractional probabilities across all strata, the estimation is indeterminate if all g terms are estimated freely. To address this, the first stratum in each row of the g terms is fixed at 0 (e.g., $g_{s,m,2,k,1} = 0$). This means that for each stock s, season m and age class a, the movement matrix (n_{strata} x n_{strata}) requires the estimation of n_{strata} parameters (n_{strata} -1 g parameters and one v parameter). The g and v parameters are assigned weak normal "priors" with mean 0 (with very low weight). Previous studies (Carruthers et al. 2011) have demonstrated that the simplified gravity modelling approach is estimable from spatial abundance indices alone, which means that the estimation will not fail for spatio-temporal strata that are sparse in terms of electronic tagging data.

For the two other movement age classes (a=1 and a=3), the g parameters and v parameters are calculated as penalized deviations from the age class 2 parameters. This allows the model to borrow information across the age classes easily when data are sparse (e.g., if data are available for age class 2 only, age classes 1 and 3 use age class 2 parameters; if age class 1 data only are available, age classes 2 and 3 borrow age class 1 parameters). For example, for age class 1:

$$g_{s,m,1,r} = g_{s,m,2,r} + \theta_{G,s,m,1,r} \tag{3.21}$$

Note that due to data sparsity it is not possible to estimate stock-specific viscosity v. It is still possible for Eastern and Western stocks to have radically different spatial distributions as determined by the g terms, but their seasonal propensity to stay in a given stratum is linked for the two stocks.

Movement parameters are estimated for individuals moving from a stratum k to a stratum r. For each stratum k, from which individuals can move, one value is assigned zero and all other possible movements are assigned an estimated parameter ψ (since rows must sum to 1, there is one less degree of freedom):

$$lnmov_{s,m,a,k,r} = \begin{cases} 0 & first\ assigned\ possible\ movement\ from\ k\ to\ r \\ \Psi_{s,m,k,r} & other\ possible\ movements\ from\ k\ to\ r \end{cases} \tag{3.22}$$

An additional movement exclusion matrix is also included in the model. This prevents any movements occurring that have not been observed for any of the three age classes of fish. This provision rules out around 1/3 of the possible seasonal tag transitions.

Compared with spatially aggregated models, initialization is more complex for spatial models, particularly those that need to accommodate seasonal movement by age and may include regional spawning and recruitment. The equilibrium unfished age structure / spatial distribution cannot be calculated analytically. For any set of model parameters, it is necessary to determine these numerically by iteratively multiplying an initial guess of age structure and spatial distribution by the movement matrix. The solution used here is to iterate the transition equations above, given a fishing mortality rate averaged over the first five years of model predictions, until the spatial distribution of stock numbers converges for each of the quarters.

Prior to this iterative process an initial guess at the spatial and age structure of stock numbers $\widehat{N}$ is made based on the movement matrix and natural mortality rate at age M:

$$\widehat{N}_{s,m,a,r} = R_{0,s} \cdot \mathrm{e}^{-\sum_1^{a-1} M_{s,a}} \cdot \sum_k \frac{1}{n_r} \cdot mov_{s,m,\mathrm{a},k,r} \tag{3.23}$$

In the years 1864 to 1964, the model does not predict catches from estimates of fishing mortality rate; instead, this historical 'spool-up' phase removes catches from the model without error. These historical catches are reconstructed for each age-class and spatio-temporal stratum. This is intended to account for meaningful landings prior to 1965 that are not accompanied by sufficient length composition data to estimate fleet selectivities in a conventional statistical catch-at-length model that is applied for the years 1965 – 2019.

Stock numbers for historical years (e.g., 1864-1964) are calculated using the same equations as model years (e.g., 1965 – 2019). The exception is that rather than using effort data, selectivities and an inverse age-length key, fishing mortality rate at age is derived from mean historical catches and the assumption is made that these are taken without error in the middle of the time step with natural mortality rate occurring both before and after fishing:

$$F_{i=1,m,a,r,f} = \begin{cases} -\log\left(1 - \frac{\bar{C}_{m,a,r,f}}{\widehat{N}_{s,m,a,r} e^{-(t_m M_{s,a})/2}}\right) & i = 1 \\ -\log\left(1 - \frac{\bar{C}_{m,a,r,f}}{\vec{N}_{s,y-1,n_m,a,r}\, e^{-(t_m M_{s,a})/2}}\right) & i > 1, m = 1 \\ -\log\left(1 - \frac{\bar{C}_{m,a,r,f}}{\vec{N}_{s,y,m-1,a,r} e^{-(t_m M_{s,a})/2}}\right) & i > 1, m > 1 \end{cases} \tag{3.24}$$

where i=1 is the first year and calculates fishing mortality rates from asymptotic numbers $\widehat{N}$.

Baseline

- Beverton-Holt with fixed steepness (see Section 9.1 for a detailed account of the stock-specific recruitment assumptions).
- Recruitment calculated from stock-wide SSB. Recruits are subsequently placed in the MED strata (Eastern stock) or in the GOM or WATL strata in proportion to the relative SSB in each of those strata (Western stock).
- Gravity movement model used to calculate a Markov movement matrix by quarter, stock and age class (e.g., Carruthers et al. 2011).

Alternative options

- Recruitment calculated from spawning strata SSB
- Markov movement matrix by quarter and stock (note: the gravity model chosen for the Baseline is a specific case of the more general Markov model).

3.3 Fleet structure and exploitation history

Table 3.1. Fishing fleets included in the operating model, based on the selectivities of fleets active historically in the Atlantic. The data described are those used in the conditioning. Catch and length composition by fleet are prepared by year, quarter, and strata from the revised CATDIS (Kimoto et al., 2021) The columns of "Strata" and "Quarter" list the strata and quarters that have catches in the revised CATDIS.

No.	Name	Gear	Flag	Strata	Quarter	Start-End*	Selectivity type/Bounds on fleet selectivity**
1	LLOTH	LL	All except Japan	All	All	1964-2019	DN; 12.5 – 412.5
2	LLJPNold	LL	Japan	All	All	1964-2009	DN; 12.5 – 362.5
3	BBold	BB	EU-Spain, EU-France	Bay of Biscay (EATL)	2,3,4	1960-2006	DN; 12.5 – 237.5
4	BBnew	BB	EU-Spain, EU-France	Bay of Biscay (EATL)	2,3,4	2007-2019	DN; 12.5 – 287.5
5	PSMEDold	PS	All except EU-Croatia	MED	1,3,4	1960-2008	DN; 12.5 – 362.5
6	PSMEDoldQ2	PS	All except EU-Croatia	MED	2	1960-2008	DN; 12.5 – 312.5
7	PSMEDnew	PS	All except EU-Croatia	MED	All	2009-2019	DN; 12.5 – 362.5
8	PSNOR	PS	Norway	NATL, EATL	3,4	1964-2019	DN; 112.5 – 337.5
9	PSHRV	PS	EU-Croatia	MED	All	1991-2019	DN; 12.5 – 287.5.
10	PSWold	PS	USA, Canada	ATW	2,3,4	1964-1984	DN; 12.5 – 337.5
11	PSWnew	PS	USA, Canada	ATW	All	1985-2019	DN; 62.5 – 312.5
12	TPold	TP	EU.Spain, Morocco, EU- Portugal	St. Gibraltar (SATL, MED)	All	1964-2011	DN; 37.5 – 337.5
13	TPnew	TP	EU-Spain, Morocco, EU- Portugal	St. Gibraltar (SATL)	2,3,4	2012-2019	DN; 37.5 – 362.5
14	RRCAN	RR	Canada	ATW, GSL	All	1964-2016	Logistic; 12.5 – 362.5
15	RRUSAFS	RR	USA	ATW	2,3,4	1964-2019	DN; 12.5 – 162.5
16	RRUSAFB	RR	USA	ATW	2,3,4	1964-2019	DN; 62.5 – 362.5
17	OTH	other	other	All	All	1964-2019	DN; 12.5 – 362.5
18	LLJPNnew	LL	Japan	WATL, SATL, NATL, EATL	All	2010-2019	DN; 62.5 – 312.5

** Catch data is used up to 2019, but the catch at size (i.e. for 2017 and later) is not updated for the purpose for reconditioning the OMs (Anon., 2021).*
*** Selectivity type DN means double normal. Bounds are the middle point in a length bin (width of length bin is 25cm) that are the lowest and highest for which lengths have been observed for each fleet.*

Baseline

An 18-fleet model based on the definitions of **Table 3.1**.

Alternative options

A proposal for alternatives may need to be developed and reviewed in the future.

4 MANAGEMENT OPTIONS

Notes:

a) The following section is included to provide some suggestions on possible structures to Candidate MP (CMP) developers of management options to be included in the CMPs. The suggestions offered are illustrative – clearly they will need to be discussed with stakeholders as the process develops.

b) As above, for convenience they have been set out in Baseline and alternative option form. It is recommended that many of the choices for the final CMP options be made later in the process, so that they can be informed by results from trials which show the pro/con trade-offs amongst such options.

c) The specifics of future CMPs will be left to their developers to determine based on the results of their application to the finalised trials. However, those candidates need to take account of the broad desired characteristics/limitations set out below.

d) HCRs need not explicitly include reference points

e) In March 2019 Panel 2 (Anon. 2019) met and began discussing their recommendations on what their objectives would be for the MSE. They also provided some guidance on preferences for some management options. This advice will be incorporated below where applicable.

4.1 Spatial strata for which TACs are set

<u>Baseline</u>

Conventional West and East/Mediterranean regions (**Figure 1.1**):

- West area: strata 1-3 (GOM, WATL, GSL)
- East area: strata 4-7 (SATL, NATL, EATL, MED)

<u>Alternative options</u>

Various possibilities exist. For example, separating out central Atlantic strata.

- A more complex 10 strata option could separate both the central Atlantic and the Caribbean (CAR).
- West: strata 1-4 (GOM, CAR, WATL, GSL).
- East+Med: strata 5-10 (SCATL, NCATL, NEATL, EATL, SEATL, MED).

However, it is suggested that consideration of such more complex options be postponed to a "second round".

4.2 Management period length for the setting of TACs

The management period is the number of years a TAC is set before the management procedure is used again to calculate a new TAC. The length of the management period must be set when implementing a CMP, and managers should be consulted on desirable management period lengths to make certain the period length is suitable for other management actions needed beyond TAC setting (e.g., fleet allocation planning, consultations, etc.). Panel 2 has indicated they might prefer to see a three-year management cycle, similar to what is currently used in Bluefin Tuna management plans (Anon. 2019). Currently CMPs are evaluated based on a two-year cycle, commencing in 2023 (see **Table App.1.1**).

<u>Baseline</u>

Every two years both a West area TAC and an East+Med area TAC are set.

<u>Alternative options</u>

i) Every three years
ii) Every four years

4.3 Upper limits on TACs

The "upper limits on TAC" allows CMP developers to put restrictions on the maximum level the TAC can achieve in the running of the CMPs.

Baseline

No upper limit

Alternative options

West	e.g. 5 000, 6 000 tons
East +Med	e.g. 30 000, 40 000 tons

4.4 Minimum extent of TAC change

The "minimum extent of TAC change" allows the CMP developer to avoid having small changes in TAC between management periods by setting a value and implementing the TAC change only if the extent of the change is at least equal to the set value. Managers might find this desirable to avoid having "trivially" small increases or decreases being incorporated in management recommendations. This restriction should be used only if it is requested by managers; otherwise it should be kept at no minimum as is the case in the Baseline below.

Baseline

No minimum (however CMPs may impose constraints at the discretion of developers).

Alternative options

West	e.g. 200, 300 tons
East +Med	e.g. 1 000, 2 000 tons

4.5 Maximum extent of TAC change

The "maximum extent of TAC change" allows CMP developers to limit the maximum allowed increase or decrease in TAC between management periods. This may help to achieve TAC stability between consecutive management periods. CMP developers can also incorporate a "maximum extent of TAC change" in the actual design of their CMP, so there are two ways to incorporate this type of constriction. The Panel 2 has provided several values of maximum extent of TAC change they would like to see (Anon 2019). These values are 20%, 30%, 40%, and outcomes where no restriction in TAC is implemented.

Baseline

West	No restriction (CMPs may impose constraints at the discretion of developers)
East +Med	No restriction (CMPs may impose constraints at the discretion of developers)

Alternative options

West	20%, 30%, 40%
East +Med	20%, 30%, 40%

Note that developers of candidate CMPs should consider including options which:

a) Override such restrictions on the maximum extent of TAC reduction if abundance indices drop below specified thresholds.

b) Allow for greater TAC increases (in terms of tonnage) if a TAC has had to be reduced to a low level and indices confirm subsequent recovery.

4.6 Technical measures

No "technical measures" are currently being implemented in the MSE. Size restrictions might be considered for a fleet and/or spatial stratum basis. However, for a "first round" it is suggested that these not be included explicitly, but instead be considered to be implemented implicitly through the selectivity prescriptions for future catches by the various fleets, which are set out under Section 6 below.

5 RECRUITMENT AND SPATIAL DISTRIBUTION SCENARIOS IN THE OPERATING MODEL

See also Section 9 of this document for additional detail on specified trials.

Recruitment deviations are estimated in two-year time blocks (i.e., the same deviation in the two years). This is necessary because the model is fitted to length-composition data (without age composition data). Due to variability in growth there are multiple age classes in each length class and, therefore, adjacent cohorts have poorly informed relative strengths. In traditional statistical-catch-at-length models this feature often leads to strong negative correlation between adjacent years in estimates of annual recruitment deviations, a poorly defined estimation problem and numerical instability of parameter estimation.

5.1 Recruitment by Stock

Western stock

Functional forms fitted to assessment outputs for the years 1965+

a) Beverton-Holt with steepness h fixed to 0.6 until 1974, then h fixed to 0.9 as of 1975. Two values of R_0 (unfished recruitment) are estimated, one for each time period (mimics the hockey-stick approach after 1975 used in past assessments).
b) Beverton-Holt with steepness h fixed to 0.6. A single value of R_0 is estimated.

Eastern stock

Functional forms fitted to assessment outputs for the years 1965+

a) Beverton-Holt with h = 0.98. Two values of R_0 (unfished recruitment) are estimated, for the periods 1950-1987 and 1988+ respectively.
b) Beverton-Holt with steepness h fixed to 0.7. A single value of R_0 is estimated (to include scenarios where recruitment overfishing could occur to test a CMP's ability to react adequately to this).

Note that, for the Eastern stock, 1965-1987 represents "low" recruitment and 1988+ "high" recruitment.

5.2 Future regime shifts

Western stock

a) None
b) After 10 years of projection, there is a change back to the pre-1974 stock-recruitment relationship (applicable only to OMs with a shift in 1975).

Eastern stock

a) None
b) After 10 years there is a change back to the 1965-1987 relationship (applicable only to OMs with a shift in 1988).

5.3 Statistical properties of recruitment deviations

Baseline

For historical years, log recruitment deviations ε_R (also called "residuals") are estimated in two-year time blocks t, starting from a lognormal prior with standard deviation σ_R = 0.354. This value is calculated from an annual value of 0.5 (a common value obtained from the RAM legacy database) using the equation for standard error: $0.354 = 0.5/2^{0.5}$.

Treating these blocks as contiguous (similarly to adjacent years for the indices below) estimates of autocorrelation $\rho_{R,2}$ and standard deviation $\sigma_{R,2}$ can be estimated for each stock s, post model conditioning (not within model fit) – this is for greater numerical stability:

$$\rho_{R,s,2} = \frac{\sum_1^{n_t-1} \varepsilon_{R,s,t,}\varepsilon_{R,s,t+1}}{\sum_1^{n_t} \varepsilon_{R,s,t}{}^2} \tag{5.1}$$

The standard deviation σ_R is calculated by:

$$\sigma_{R,s,2} = \sqrt{\frac{1}{n_t-1}\sum_1^{n_t}\left(\varepsilon_{R,s,t} - \bar{\varepsilon}_{R,s}\right)^2} \tag{5.2}$$

For future projection years, annual recruitment deviations (1-year blocks) for each stock are simulated from lognormal distributions with variance and autocorrelation derived from the historical residuals for that stock. It is necessary therefore that the estimated statistical properties for 2-year blocks (autocorrelation $\rho_{R,s,2}$ and standard deviation $\sigma_{R,s,2}$) are converted to annual equivalents ($\rho_{R,s,1}$ and $\sigma_{R,s,1}$).

For future projection years and for each stock separately, the annual variance in recruitment deviations $\sigma^2_{R,s,1}$ and annual first order autocorrelation in recruitment deviations $\rho_{R,s,1}$ can be derived analytically from 2-year blocked estimates of variance $\sigma^2_{R,s,2}$ and 'lag-1' $\rho_{R,s,2}$ autocorrelation by inverting the relationships:

$$\sigma^2_{R,s,2} = \sigma^2_{R,s,1}\left(1 + \rho_{R,s,1}\right)/2 \tag{5.3}$$

$$\rho_{R,s,2} = \rho_{R,s,1}\left(1 + \rho_{R,s,1}\right)/2 \tag{5.4}$$

It follows that $\sigma_{R,s,1}$ is calculated by:

$$\sigma^2_{R,s,1} = 2\sigma^2_{R,s,2}/\left(1 + \rho_{R,s,1}\right) \tag{5.5}$$

If $\rho_{R,s,2}$ is positive, $\rho_{R,s,1}$ is given by:

$$\rho_{R,s,1} = \frac{\left(-1+\sqrt{1+8\cdot\rho_{R,s,2}}\right)}{2} \tag{5.6}$$

Since equation 5.4 is quadratic, there are two solutions for $\rho_{R,1}$ given a negative value of $\rho_{R,2}$ (**Figure 5.1**).

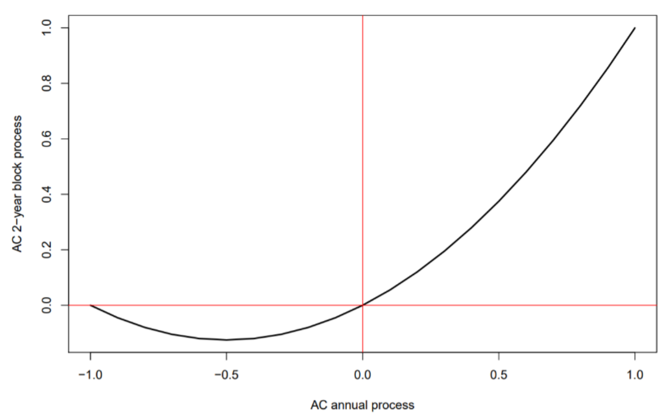

Figure 5.1. $p_{R,s,2}$ as a function of $p_{R,s,1}$ (courtesy of C. Fernandez)

In such cases, the larger value of $\rho_{R,s,1}$ (i.e., the one of smaller absolute magnitude) is used. If $\rho_{R,s,2}$ is less than -0.125, there is no solution to equation 5.4, so the value corresponding to the lower possible value of $\rho_{R,s,2}$, i.e., $\rho_{R,s,1}$ = -0.5, is used. Note that this can occur because of higher lag AC effects; the model used here assumes that there are no AC effects beyond a 1-year lag.

For the simulation of future recruitment, stochastic residuals are generated by first sampling uncorrelated normal recruitment residuals for each projection year y and simulation j:

$$\varepsilon_{R,U,s,y,j} \sim N\left(0, \sigma^2_{R,s,1}\right) \tag{5.6}$$

For the first projection year immediately after the last year of model conditioning the autocorrelated residual error is calculated from the last 2-year-block residual in time T:

$$\varepsilon_{R,s,y,j} = \rho_{R,s,1}\varepsilon_{R,s,T} + \sqrt{1-\rho_{R,s,1}{}^2}\,\varepsilon_{R,U,s,y,j} - (1-\rho_{R,s,1})(\sigma^2_{R,s,1})/2 \tag{5.7}$$

For all subsequent projection years autocorrelated residuals are calculated by:

$$\varepsilon_{R,s,y,j} = \rho_{R,s,1}\varepsilon_{R,s,y-1,j} + \sqrt{1-\rho_{R,s,1}{}^2}\,\varepsilon_{R,U,s,y,j} - (1-\rho_{R,s,1})(\sigma^2_{R,s,1})/2 \tag{5.8}$$

5.4 Possible future spatial distributional changes (movement)

Plausible options for future distributional changes (in relative terms) in response to changes in abundance and to possible environmental changes will be considered in a "second round".

6 FUTURE CATCHES

6.1 Catches between 2020 and 2022

Baseline

Table 6.1. Recent allocations (tons) by fleet

No	Fleet	Area	Country	2020	2021*	2022*
1	LLOTH	Med	all except Japan	1823.633	1766.834	1766.834
1	LLOTH	East	all except Japan	539.6648	660.2877	660.2877
1	LLOTH	West	all except Japan	175.9687	231.234	268.2315
18	LLJPN	East	Japan	2781.631	2781.631	2781.631
18	LLJPN	West	Japan	407.5781	407.48	664.52
4	BBnew	East	France and Spain in Bay of Biscay	135.2339	193.9202	193.9202
7	PSMEDnew	Med	All PS except Croatia in Med	20047.96	21149.66	21149.66
8	PSNOR	Med	Norway	189.482	290.8452	290.8452
9	PSHRV	Med	Croatia	829.0691	870.0258	870.0258
11	PSWnew	West	USA and Canada	0	0	0
13	TPnew	East	Spain, Morocco and Portugal	5885.417	5484.5	5484.5
13	TPnew	Med	Spain and Morocco	5.6204	3.013779	3.013779
14	RRCAN	West	Canada	545.6398	425.0523	447.531
15	RRUSAFS	West	USA	192.554	200.5365	211.2939
16	RRUSAFB	West	USA	933.805	991.3908	1044.572
17	OTHERS	Med	All others	1458.068	1367.136	1367.136
17	OTHERS	East	All others	1342.382	1432.144	1432.144
17	OTHERS	West	All others	13.816	94.306	89.852

** The 2021 and 2022 catch assumptions are based on the 2021 and 2022 TACs for the East the West areas.*

a) Future catches will be taken to equal future TACs (up to a maximum harvest proportion $U_{max} = 90\%$ in each stratum-quarter).
b) The catch by fleet in 2017 was calculated from ICCAT Task 1 (**Table 6.1**).
c) The allocation of future catches amongst fleets will be set equal to the average decided by the Commission for the period 2018-2020 (**Table 6.1**).
d) The spatial distribution (see Section 1) of these future catches will be set equal to the average over 2017-2019 (last three years of model-estimated spatio-temporal catch distribution).
e) The selectivity function for each fleet for the most recent period for which this is estimated in the conditioning of the operating model in the trial concerned will be taken to apply for all future years.

f) TAC and catches are fixed into projection model (2019 and after) based on TACs for 2019-2021 and TACs as reflected in the recommendations: [Rec. 17-06], [Rec.17-07], and [Rec. 18-02], with the West area TAC for 2022 (still to be finalised by the Commission) set equal to the 2021 value (for catch reconstruction methods see Kimoto et al. 2021).

Alternative options

Clearly many are possible, but are probably best delayed until a "second round". Were substantial changes to eventuate during a period when a CMP was in operation, this would in any case likely necessitate re-tuning and re-testing, or a modified CMP.

The impacts of possible IUU catches should perhaps be considered under robustness trials (see Section 9 below).

6.2 Equations for calculating projected catch in numbers by stock, fleet, spatial strata, year and quarter based on TACs specified in weight by management area

Management procedures provide TAC advice in weight by East and West management area. In order to project the simulated stocks accounting for feedback from these TAC recommendations they must be converted to an estimate of fishing mortality rate at age for each fleet (e.g., LLJPN), spatial stratum (e.g., West Atlantic), year and quarter. These catches cannot exceed the maximum available number of fish in any spatial stratum at any time, so that they must be allocated across fleets in a plausible way; the TACs must sum to the management area totals by weight and optionally, so that there should be the ability for missed quota in one fleet to be reallocated (traded) to other fleets where catches may still be possible.

Under MSE projections (after 2017), TACs by East-West management area are distributed among fleets according to a management-area-specific fraction of the TAC among fleets *A,* a fleet specific seasonal and spatial fraction *D*, and the predicted age composition of catches $\hat{V}_{s,y,m,a,r,f}$

$$C_{s,y,m,a,r,f} = \begin{cases} A_{f,west} \cdot D_{f,m,r} \cdot TAC_{y,west} \cdot \hat{V}_{s,y,m,a,r,f,west} & r \leq 3 \\ A_{f,east} \cdot D_{f,m,r} \cdot TAC_{y,east} \cdot \hat{V}_{s,y,m,a,r,f,east} & r \geq 4 \end{cases} \tag{6.1}$$

where *west* is the West management area, *east* is the East management area, *A* are fractions (sum to 1 for each management area) of the 2020 allocations (**Table 6.1**), *D* terms are fractions (sum to 1 for each fleet) calculated from the model predicted fishing mortality rates over the last three historical years (2015-2017), and

$$\hat{V}_{s,y,m,a,r,f,west} = \frac{\vec{N}_{s,y,m,a,r} \cdot s_{f,a}}{\sum_m \sum_a \sum_{r=1}^{3} \vec{N}_{s,y,m,a,r} \cdot s_{f,a}} \tag{6.2}$$

$$\hat{V}_{s,y,m,a,r,f,east} = \frac{\vec{N}_{s,y,m,a,r} \cdot s_{f,a}}{\sum_m \sum_a \sum_{r=4}^{7} \vec{N}_{s,y,m,a,r} \cdot s_{f,a}} \tag{6.3}$$

where $S_{f,a}$ is the fleet selectivity at age.

Equations 6.1 – 6.3 keep both the allocations among fleets and the spatio-temporal distribution of catches by fleet constant, whilst allowing for varying catch at age in response to availability of numbers at age. Dynamic spatio-temporal fishing distribution (*D* by projection year) would require a fleet dynamics model that has not been developed for Atlantic bluefin tuna.

It is possible for MPs to prescribe catches that are higher than the available stock biomass (are not possible). In these cases, the harvest rate *U* exceeds 1. *U* is calculated as the predicted catch weight for the fleet C_w divided by the total biomass *B* across all fleets in that year, quarter and spatial stratum:

$$U_{y,m,r,f} = C_{W,y,m,r,f} / B_{y,m,r,f} \tag{6.4}$$

where catch weight C_W is calculated by the sum product of the catches at age (equation 6.1) and weight at age.

To avoid catches exceeding available biomass a somewhat arbitrary maximum harvest rate U_{max} is used (the default value is 90%).

In cases where $U_{y,m,r,f} > U_{max}$ there are two options that have currently been implemented in the ABTMSE R framework: (1) remove these catches and (2) reallocate these catches. Reallocation is selected as the default setting (it can be switched off) in order to approximate quota trading among fleets which may be expected to occur when quotas cannot be caught by one or more fleet.

The reallocation rule redistributes catches in excess of the maximum harvest rate C_{Umax}, across all fleet-spatial strata (f, r) (in the same management area) that are below the U_{max} constraint:

$$C_{Umax,y,m,r,f} = \begin{cases} C_{W,y,m,r,f} \cdot \left(U_{y,m,r,f} - U_{max}\right)/U_{y,m,r,f} & U_{y,m,r,f} \geq U_{max} \\ 0 & U_{y,m,r,f} < U_{max} \end{cases} \tag{6.5}$$

$$\vec{C}_{W,y,m,r,f} = \begin{cases} C_{W,y,m,r,f} - C_{Umax,y,m,r,f} & U_{y,m,r,f} \geq U_{max} \\ C_{W,y,m,r,f} + C_{D,y,m,r,f} \cdot \sum_f C_{Umax,y,m,r,f} & U_{y,m,r,f} < U_{max} \end{cases} \tag{6.6}$$

where the reallocated catches by weight $\vec{C}_W$ either remove catches above U_{max} (where $U > U_{max}$) or redistribute the total C_{Umax} in proportion to the catches of fleets below the U_{max} constraint C_D given by:

$$C_{D,y,m,r,f} = \frac{\theta_{C,y,m,r,f} C_{W,y,m,r,f}}{\sum_f \theta_{C,y,m,r,f} C_{W,y,m,r,f}} \tag{6.7}$$

where θ_C is 1 where $U < U_{max}$ and 0 otherwise.

Note that when all fleets exceed the U_{max} constraint, equation 6.6. ensures that in all cases, catches above the constraint are removed without reallocation. This reallocation rule is simple, numerically robust and does not require additional fleet dynamics models. The relatively fine temporal resolution of the operating model (quarterly) means that although it is theoretically possible for C_{Umax} catches to be reallocated to other fleets such that the new catches $\vec{C}_W$ then exceed U_{max}, these catch reallocations are generally very small and occur very progressively; consequently, this rarely occurs.

7 GENERATION OF FUTURE DATA FOR INPUT TO CANDIDATE MANAGEMENT PROCEDURES

Note that these are for use as input to CMPs, so need to be chosen carefully from a set of those highly likely to be regularly (i.e., annually) available. This is because the application of a CMP relies on these data being available in this way, so difficulties can (and have in other cases) obviously arise should they fail to do so. Though any CMP proposed should include a rule to deal with the absence of just one future value from an input series, any more than that would require re-tuning and re-testing of a modified CMP, and ideally this is to be avoided given the extra associated costs.

Consideration is also needed of the "delays" associated in such data becoming available for input to an CMP. When a TAC is set for year *y*, the last year of finalised data at the time of setting the TAC is *y*-2 for surveys and CPUE indices and *y*-3 for catch data. For years *y*-2 and *y*-1 the catch is assumed to be equal to the TAC.

- TAC implementation year = *y*
- Commission decision year = *y*-1
- SCRS advice year = *y*-1
- CPUE/Independent last data year = *y*-2

Therefore CPUE/independent data would have to be finalized up until year *y*-2 and provided to SCRS meeting that takes place in Sept of year *y*-1.

In the closed loop projections of the ABTMSE package, the most recent (last available) index observation of the simulated dataset (dset) is *y*-2, and the most recent (last available) catch observation is *y*-1.

7.1 Indices

Baseline

Indices are simulated in projections based on the operating model-specific fit to the indices (or values specified during the 2021 April BFT intersessional meeting: Anon. 2021), including lognormal error (STD) and lag-1 autocorrelation in residuals (AC) (**Table 7.1**)

Table 7.1. Index selection and simulation for potential inclusion in CMPs

Index	Details	Selectivity	Recommended for CMPs	STD value*	AC*
CAN GSL RR	1988-2020, Q3, GSL	14: RRCAN	No	-	-
CAN SWNS RR	1996-2020, Q3, W Atl	14: RRCAN	Yes	OM-estim	OM-estim
US RR 66-144	1995-2020, Q3, W Atl	15: RRUSAFS (50 –150cm)	Yes	OM-estim	OM-estim
US RR 66-114	1995-2020, Q3, W Atl	15: RRUSAFS (50 –125cm)	No***	OM-estim	OM-estim
US RR 115-144	1995-2021, Q3, W Atl	15: RRUSAFS (100 – 150cm)	No***	OM-estim	OM-estim
US RR 177+	1993-2020, Q3, W Atl	16: RRUSAFB (175cm+)	No	-	-
JPN LL West2	2010-2020, Q4, W Atl	18: LLJPNnew	Yes	OM-estim	OM-estim
US-MEX GOM PLL	1994-2019, Q2, GOM	1: LLOTH	Yes	OM-estim	OM-estim
GOM LAR SUV	1977-2021 (gaps 1979-1980, 1985, 2020), Q2, GOM	SSB	Yes	OM-estim	OM-estim
CAN ACO SUV2	2017-2018, Q3, GSL	14: RRCAN (150cm+)	No**	-	-
MOR POR TRAP	2012-2020, Q2, S Atl	13: TPnew	Yes	OM-estim	OM-estim
JPN LL NEAtl2	2010-2019, Q4, N Atl	18: LLJPNnew	Yes	OM-estim	OM-estim
FR AER SUV2	2009-2021 (gap 2013), Q3, Med	15: RRUSAFS	Yes	OM-estim	OM-estim
GBYP AER SUV BAR	2010-2018 (gaps 2012, 2014, 2016), Q2, Med	SSB	Yes	0.45#	0.2#
W-MED LAR SUV	2001-2019 (gaps 2006, 2007, 2009, 2011), Q2, Med	SSB	Yes	OM-estim (years 2012-2019)	OM-estim (years 2012-2019)

** OM-estim means OM-specific estimates from the index residuals of the corresponding OM fit (Section 7.5). When the estimated AC is < 0, it is fixed at AC=0 for the projections with that OM.*
*** The Canadian acoustic survey index is simulated in the BFT MSE package, but should not be used in CMPs at this time because of uncertainty about calibration in the change to a different vessel.*
**** Not recommended for CMPs but still projected for sensitivity runs.*
 # GBYP AER SUV BAR index will be refit by contractor and SE and AC re-evaluated with a preference given to using estimated SE and AC values

These generated indices must maintain the same methods for their construction in future years; changes to how the indices are constructed would not be allowed during an accepted MP's period of use.

Some CMPs may use annual catch (removals) observations in addition to the simulated indices. For the Baseline, simulated annual catch data are assumed to have been observed with error and a log-normal CV of 2.5% (95% of observations are within +/- 5% of the true catch that was taken).

While not all of the indices are being used for projections, this does not imply that they should be discontinued or not updated and reviewed by the SCRS BFT species group. It will also be important to have these updated indices for model re-conditioning when the MSE is re-run (which would be conducted after a set interval to be determined by the Commission).

<u>Alternative options</u>

Catch-at-length series could also be considered for inclusion as CMP inputs, but raise further technical complications regarding the specification of how they are generated, so have been deferred from consideration until a "second round".

A 'perfect information' observation error model (suitable for CMP testing) that includes essentially no observation error or autocorrelation in indices, or observation error in catches.

A 'bad' observation error model that is the same as the Baseline but includes the estimated non-linearity in indices with biomass, and a 10% lognormal CV in annual catch data.

7.2 Relationships with abundance

For Reference Grid trials, abundance indices are taken to be linearly proportional to the appropriate component of the underlying model biomass in the spatial stratum and quarter concerned. Possible alternatives to this are considered under Robustness trials (Section 9.2). In robustness trials where hyperstability is included, the *beta* parameter controls how quickly the index I responds to vulnerable biomass B: $I = qB^{beta}$. This is set to a value lower than 1 (e.g., 0.5) and the operating models are refitted assuming this value; this same value is then also assumed when generating indices in projection years.

7.3 Statistical properties

<u>Baseline</u>

Residuals are taken to be log normally distributed; The standard deviation and lag-1 autocorrelation of the log residual error (specified in **Table 7.1**) are invariant over time.

<u>Alternative options</u>

a) Fix σ values for all trials based on a central trial from the Reference Grid (Section 9.1).
b) If additional CPUE indices to those initially suggested are included, residuals need to be examined for correlation, with this being taken into account in generating future values.
c) In a "second round", take correlations amongst indices into account.

7.4 Other aspects

Note that consideration should at some stage also be given to new data types that are only now becoming available (e.g., genetic tagging). These will not at this stage have been collected over a sufficient length of time to be able to serve as CMP inputs, but the overall CMP testing process can be used to provide insight into their potential future utility.

7.5 Equations for the simulation of indices

The fishery dependent CPUE indices and some of the fishery independent indices identified in **Table 7.1** correspond to the length selectivity of a fleet and are assumed to correspond to either vulnerable numbers or vulnerable biomass. Most of the fishery-independent indices are assumed to correspond to spawning stock biomass. For all indices, the observed index and the model predicted index are specific to a particular quarter and spatial stratum (**Tables 2.1** and **2.2**). To simplify the algebra of this section - which focuses on simulating indices in the future according to their statistical properties - equations are provided for a single index and its corresponding quarter and spatial stratum (represented by a single subscript i). For a more detailed breakdown of the calculation of model predicted indices see Section 8.

When an operating model is fitted, a maximum likelihood estimate of the model predicted index $\hat{I}$, is calculated for historical years where there are observations I. The log residual is assumed to be normally distributed and is calculated by:

$$\varepsilon_{i,y} = ln(I_{i,y}) - ln(\hat{I}_{i,y}) \tag{7.1}$$

For annual indices (without gaps), the first order (lag-1) autocorrelation of residuals ρ is calculated by:

$$\rho_i = \frac{\sum_1^{n_y-1} \varepsilon_{i,y}, \varepsilon_{i,y+1}}{\sum_1^{n_y} \varepsilon_{i,y}{}^2} \tag{7.2}$$

The standard deviation σ is calculated by:

$$\sigma_i = \sqrt{\frac{1}{n_y-1} \sum_1^{n_y} \left(\varepsilon_{i,y} - \bar{\varepsilon_i}\right)^2} \tag{7.3}$$

After calculating these statistics based on the fit to historical data, the future errors must be simulated for each index. There are two situations that must be covered: (1) no observations of the index after the period of model fitting; and (2) observations of the index that have occurred after model fitting and whose values are already known.

The first situation is the simplest. For a given simulation *j*, for all years after the last year of model fitting, uncorrelated residuals are drawn from a normal distribution with mean 0 (no bias correction was applied in model fitting):

$$\varepsilon_{U,i,y,j} \sim N(0, \sigma_i^2) \tag{7.4}$$

For the first projection year immediately after the last year of model conditioning, the autocorrelated residual error is calculated by:

$$\varepsilon_{i,y,j} = \rho_i \varepsilon_{i,y-1} + \sqrt{1-\rho_i{}^2}\, \varepsilon_{U,i,y,j} - (1-\rho_i)(\sigma_i^2)/2 \tag{7.5}$$

For all subsequent projection years autocorrelated residuals are calculated by:

$$\varepsilon_{i,y,j} = \rho_i \varepsilon_{i,y-1,j} + \sqrt{1-\rho_i{}^2}\, \varepsilon_{U,i,y,j} - (1-\rho_i)(\sigma_i^2)/2 \tag{7.6}$$

In the second situation where index observations have occurred after model fitting, residuals must be calculated for those observations in order to calculate projected autocorrelation in residuals correctly. Since the residual error of these observations is now based on projected biomass subject to simulation-specific process errors, these residuals are calculated specific to each simulation *j*:

$$\varepsilon_{i,y,j} = ln(I_{i,y}) - ln(\hat{I}_{i,y,j}) \tag{7.7}$$

Then, for subsequent projection years after the last observation (2019 or before), error terms are calculated from equation 7.6.

7.6 Stochastic versus deterministic projections

The ABTMSE R package includes both stochastic and deterministic operating models (OMs).

The deterministic OMs differ from stochastic OMs in that they do not include stochasticity in recruitment or in the generation of error in observations of catches and indices. Since they do not vary for a specific OM, there is only one simulation in each deterministic OM (the stochastic OMs contain 48 simulations). The role of the deterministic OMs is to provide a rapid check of candidate management procedure (CMP) performance, on the ground that a CMP that does not perform well under deterministic conditions is unlikely to perform well in stochastic projections.

The stochastic OMs are intended to be the primary basis for the calculation and presentation of CMP performance.

8 PARAMETERS AND CONDITIONING OF OPERATING MODELS

8.1 Fixed parameters

Table 8.1. The parameters that are fixed (user specified).

Parameter	Number of parameters	Symbol
Steepness	$\geq n_{stocks}$	*H*
Maximum length	n_{stocks}	*Linf*
Growth rate	n_{stocks}	*K*
Age at length zero	n_{stocks}	t_0
Natural mortality rate at age	$n_{ages} \cdot n_{stocks}$	*M*
Selectivity of at least one fleet	*2-3*	*Θ*
Maturity at age	$n_{ages} \cdot n_{stocks}$	*Mat*

Table 8.2. Parameter values of Reference Grid OMs.

Parameter	Western stock	Eastern stock	
Steepness	0.6 changing to 0.9 in 1975	0.98	
(Bev. -Holt)	0.6	0.7	
Len-Wt a (Rodriquez-Marin et al. 2015)	1.77054 E-5	3.50801 E-5	
Len-Wt b	3.001252	2.878451	
Type	***Richards growth*** *(Allioud et al., 2017)*	***von Bert. Growth*** *(Cort, et al. 2014)*	
A2	34		
***L1* (cm)**	33.0		
***L2* (cm)**	270.6	***Linf* (cm)**	318.85
K	0.22	***K***	0.093
p_0	-0.12	***t0***	-0.97
a_{CV} (*CVlen = a L +b*) (Allioud et al., 2017)	0.06	0.06	
b_{CV}	5.84	5.84	

Natural mortality rate at age (Eastern and Western)

Age	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11-13**	**14**	**15-25**	**26+**
High	0.38	0.30	0.24	0.20	0.18	0.16	0.14	0.13	0.12	0.12	0.11	0.10	0.10	0.10
Low	0.36	0.27	0.21	0.17	0.14	0.12	0.11	0.10	0.09	0.09	0.08	0.08	0.07	0.47

Selectivity of at least one fleet	Fleet#14 'CAN RR' is assumed to be logistic (**Table 3.1**)

Spawning fraction

Age	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13+**
Younger	0	0	0.25	0.5	1	1	1	1	1	1	1	1	1
Older (East)	0	0	0.15	0.30	0.45	0.6	0.75	0.9	1	1	1	1	1
Older (West)	0	0	0	0	0	0	0.01	0.04	0.19	0.56	0.88	0.98	1

West stock mixing	1% western stock biomass in East area on average from 1965-2016

8.2 Estimated parameters

The majority of parameters estimated by the model relate to movement probabilities and annual recruitment deviations (**Table 8.3**).

Table 8.3. The parameters estimated by the model. The example is for a bluefin tuna operating model of 7 strata (**Figure 1.1**), 4 quarters, 18 fleets, 55 years, 3 movement age classes and two estimated phases of stock-recruitment per stock (e.g., OM_1, **Figure 9.2**).

Parameter	**Number of parameters**	
Unfished recruitment (recruitment level 1)	$2 \cdot n_{stocks}$	4
Length at modal selectivity	n_{fleets}	18
Ascending precision of selectivity	n_{fleets}	18
Descending precision of selectivity	n_{fleets} -1	17
Recruitment deviations	$(n_{years}$ / 2 - 2) $\cdot n_{stocks}$	54
Fleet catchability (q)	n_{fleets}	18
F deviation (FD)	$n_{quarters} \cdot n_{strata}$	28
Annual F (FA) deviation	$n_{quarters} \cdot n_{strata} \cdot (n_{years}$ -1)	1512
Movement	$n_{strata} \cdot n_{quarters} \cdot n_{stocks} \cdot n_{mov\text{-}ages}$	168
	Total	1837

Table 8.4. "Prior" probability distributions (penalty functions) for model parameters with mean μ and standard deviation σ, and lower and upper bounds LB and UB, respectively.

Parameter	**Prior**	**Likelihood component**
All operating models		
Unfished recruitment	log-uniform(LB = 11, UB = 16.5)	$-lnL_{rec}$
All Selectivity parameters (θ)	lognormal(μ = 0, σ = 4) (LB = -5.0, UB = 5.0)	$-lnL_{sel}$
Fishing fleet catchability (q) (mean F per fleet)	log-uniform(LB = -10.0, UB = 1.0)	$-lnL_{q}$
F deviation (FD)	lognormal(μ = 0, σ = 1)	$-lnL_{FD}$
Annual F deviation (FA)	lognormal(μ = 0, σ = 1)	$-lnL_{FA}$
Movement parameters (g, v, θ_G)	lognormal(μ = 0, σ = 4) (LB = -8.0, UB = 8.0)	$-lnL_{mov}$
Recruitment deviations (2-year blocks)	lognormal(μ = 0, σ = 0.353)	$-lnL_{recdev}$
Unfished recruitment change (applicable only to the level 1 and 3 recruitment scenarios)	lognormal(μ = 0, σ = 0.4)	$-lnL_{R0dif}$
Asymptotic western stock mixing	Lognormal($\mu_{specified}$, σ = 0.025)	$-lnL_{mix}$
Mean spawning stock biomass by area	Lognormal($\mu_{specified}$, σ = 0.05)	$-lnL_{muSSB}$
Seasonal distribution of vulnerable numbers of 125cm + fish	Lognormal($\mu_{specified}$, σ = 0.025)	$-lnL_{natal}$

A summary of likelihood functions can be found in **Table 8.5**.

For each fleet f, total predicted catches in weight $\hat{C}$, are calculated from the Baranov equation:

$$\hat{C}_{y,m,r,f} = \sum_s \sum_a w_{s,a} \cdot N_{s,y,m,a,r} \cdot (1 - e^{-Z_{s,y,m,a,r}}) \cdot \left(\frac{F_{s,y,m,a,r,f}}{Z_{s,y,m,a,r}}\right) \tag{8.1}$$

Similarly, predicted catches in numbers at age (CAA) are given by:

$$\widehat{CAA}_{s,y,m,a,r,f} = N_{s,y,m,a,r} \cdot (1 - e^{-Z_{s,y,m,a,r}}) \cdot \left(\frac{F_{y,m,a,r,f}}{Z_{s,y,m,a,r}}\right) \tag{8.2}$$

This can be converted to a prediction of total catches in numbers by length class CAL using a stock specific inverse age-length key, LAK:

$$\widehat{CAL}_{y,m,l,r,f} = \sum_s \sum_a \widehat{CAA}_{s,y,m,a,r,f} \cdot LAK_{s,a,l} \tag{8.3}$$

For fishery independent indices of spawning stock biomass, the model predicts spawning stock biomass indices $\widehat{Is}$, for a specific index *i*, and stratum *r*:

$$\widehat{Is}_{i,s,y,r} = q_i \, SSB_{s,y,r} \tag{8.4}$$

For fishery independent indices and fishery dependent indices based on vulnerable biomass (noting that for the Canadian acoustic survey this is calculated by vulnerable numbers) the model predicts exploitable biomass indices $\hat{I}$, by fleet:

$$\hat{I}_{i,y,m,r,f} = q_i \;\; V_{y,m,r,f} \tag{8.5}$$

where exploitable biomass *V* is calculated as:

$$V_{y,m,r,f} = \sum_l \left(s_{f,l} \cdot \sum_s \sum_a \left(N_{s,y,m,a,r,f} \cdot LAK_{s,a,l} \cdot w_{s,a} \right) \right) \tag{8.6}$$

The model predicts stock of origin composition of catches (fraction of Eastern origin) by movement age class ac, $\hat{R}$, from predicted catch numbers at age:

$$\hat{R}_{y,m,r,f,ac} = \begin{cases} \sum_f \sum_{a=1}^{4} \widehat{CAA}_{s=1,y,m,a,r,f} / \sum_s \sum_f \sum_{a=1}^{4} \widehat{CAA}_{s,y,m,a,r,f} & ac = 1 \\ \sum_f \sum_{a=5}^{8} \widehat{CAA}_{s=1,y,m,a,r,f} / \sum_s \sum_f \sum_{a=5}^{8} \widehat{CAA}_{s,y,m,a,r,f} & ac = 2 \\ \sum_f \sum_{a=9} \widehat{CAA}_{s=1,y,m,a,r,f} / \sum_s \sum_f \sum_{9} \widehat{CAA}_{s,y,m,a,r,f} & ac = 3 \end{cases} \tag{8.7}$$

A log-normal likelihood function (without constant terms) is assumed for total catches by fleet. The negative log-likelihood is calculated as:

$$-lnL_c = \sum_y \sum_m \sum_r \sum_f \frac{\left(\ln(\hat{C}_{y,m,r,f}) - \ln(C_{y,m,r,f})\right)^2}{\sigma^2_{catch}} \tag{8.8}$$

Total annual catches of each fleet are fitted assuming very high precision: a CV (σ_{catch}) of 1%.

Similarly, the negative log-likelihood components for indices of exploitable biomass and spawning stock biomass are calculated as:

$$-lnL_i = \sum_y \sum_m \sum_r \sum_f \frac{\left(ln(\hat{I}_{y,m,r,f}) - ln(I_{y,m,r,f})\right)^2}{\sigma^2_{i,y}} \tag{8.9}$$

and

$$-lnL_{SSB} = \sum_s \sum_y \frac{\left(ln(\widehat{Is}_{s,y}) - ln(Is_{s,y})\right)^2}{\sigma^2_{S,y}} \tag{8.10}$$

The negative log-likelihood component for length composition data is calculated (for positive observations of length composition only) by:

$$-lnL_{CAL} = -\sum_y \sum_m \sum_l \sum_r \sum_f \hat{p}_{y,m,l,r,f} \left(\ln(\hat{p}_{y,m,l,r,f}) - \ln(p_{y,m,l,r,f}) \right)^2 \tag{8.11}$$

where the model predicted fraction, $\hat{p}$, of catch numbers in each length class is calculated as:

$$\hat{p}_{y,m,l,r,f} = \widehat{CAL}_{y,m,l,r,f} / \sum_l \widehat{CAL}_{y,m,l,r,f} \tag{8.12}$$

The negative log-likelihood component for electronic tagging data of known stock of origin (SOO), released in year *y*, quarter *m*, stratum *r* and caught in the subsequent quarter in stratum *k* is calculated from a multinomial likelihood function as:

$$-lnL_{ET} = -\sum_s \sum_y \sum_m \sum_r \sum_k ET_{s,m,r} \cdot ln\left(mov_{s,m+1,r}\right) \tag{8.13}$$

The negative log-likelihood component for stock of origin data is calculated assuming a normal likelihood function (without constants) comparing $\hat{r}$ estimated from the operating model, with r derived applying a mixture model (Carruthers and Butterworth 2018) to assignment scores from genetics and otolith microchemistry data (here the subscript i is used to denote a specific year, area and age class):

$$-lnL_{SOO} = \sum_i \frac{(r_i - \hat{r}_i)^2}{{\sigma_{i,s}}^2} \tag{8.14}$$

where the operating model estimated logit fraction Eastern fish for the i^{th} strata, $\hat{r}_i$ is calculated from the operating model predicted ratio of Eastern fish in the catch $\hat{R}_i$ (see equation 8.7):

$$\hat{r}_i = ln\left(\hat{R}_i/(1-\hat{R}_i)\right).$$

For OMs in which R_0 changes in some past year, the following "prior" distribution is used for the extent of the change (see **Table 8.4**):

$$\sum_y \sum_m \sum_r \sum_f \frac{\left(\ln(R_{0,1}) - \ln(\mathrm{R}_{0,2})\right)^2}{\sigma^2_{R0dif}} \tag{8.15}$$

The "prior" was found to be required to ensure models could converge – without it there was very little data to inform the estimates of R0 in the later period. The highest possible value for the standard deviation σ^2_{R0dif} (the vaguest "prior") was chosen that could allow models to converge reliably across all reference case operating models.

The latest version of M3 (v6.6+) allows for "priors" for western stock mixing and SSB scale (in order to bracket operating model scenarios). Western mixing is defined as the mean fraction of western spawning stock biomass found in the East area, W_{mix}. The model estimates can then be compared with a specified prior level of mixing $\widehat{W}_{mix}$ via a log-likelihood function:

$$-lnL_{mix} = \frac{\left(ln(\widehat{W}_{mix}) - ln(\mathrm{W}_{mix})\right)^2}{\sigma^2_{mix}} \tag{8.16}$$

Similarly, mean spawning stock biomass by area can be specified as a "prior" $\widehat{SSB}_{mu,area}$ (1968-2015 in the West, 1974-2015 in the East – the years of the VPA assessments which are used for purposes of comparison).

$$-lnL_{SSBmu} = \frac{\left(ln(\widehat{SSB}_{mu,west}) - ln(\mathrm{SSB}_{mu,west})\right)^2}{\sigma^2_{SSBmu}} + \frac{\left(ln(\widehat{SSB}_{mu,east}) - ln(\mathrm{SSB}_{mu,east})\right)^2}{\sigma^2_{SSBmu}} \tag{8.17}$$

Additionally, the natal spawning areas of the Mediterranean and Gulf of Mexico include a seasonal penalty function that controls the fraction of vulnerable numbers in each area under equilibrium conditions.

$$-lnL_{natal} = \sum_{s=1}^{4} \frac{\left(ln(\hat{\gamma}_{s,GOM}) - ln(\gamma_{s,GOM})\right)^2}{\sigma^2_{natal}} + \sum_{s=1}^{4} \frac{\left(ln(\hat{\gamma}_{s,MED}) - ln(\gamma_{s,MED})\right)^2}{\sigma^2_{natal}} \tag{8.17a}$$

where γ and $\hat{\gamma}$ are the observed and model predicted fractions of 125cm+ fish in each quarter s, and natal area (Gulf of Mexico, 'GOM' or Mediterranean, 'MED').

The global penalised negative log-likelihood $-lnL_T$, to be minimized is the summation of the weighted negative log-likelihood components for the data and "priors" (**Table 8.4**):

$$\begin{aligned} -lnL_T = -[\ & \omega_c \cdot lnL_c + \omega_i \cdot lnL_i + \omega_{SSB} \cdot lnL_{SSB} + \omega_{CAL} \cdot lnL_{CAL} + \\ & \omega_{ET} \cdot lnL_{ET} + \omega_{SOO} \cdot lnL_{SOO} + \omega_{sel} \cdot lnL_{sel} + \omega_{FD} \cdot lnL_{FD} + \omega_{FA} \cdot lnL_{FA} + \omega_{mov} \cdot lnL_{mov} + \\ & \omega_{recdev} \cdot lnL_{recdev} + \omega_{R0dif} \cdot lnL_{R0dif} + \omega_{mix} \cdot lnL_{mix} + \omega_{SSBmu} \cdot lnL_{SSBmu} + \omega_{natal} \cdot \\ & lnL_{natal}\ ] \end{aligned} \tag{8.18}$$

Table 8.5. Summary of the negative log-likelihood function contributions from various data

Type of data	Disaggregation	Function	Likelihood component
Total catches (weight)	year, quarter, stratum, fleet	Log-normal	lnL_c
Index of exploitable biomass (assessment CPUE index)	year, quarter, stratum, fleet	Log-normal	$lnLi$
Index of spawning stock biomass (e.g., a larval survey)	year, quarter, stratum, stock	Log-normal	lnL_{SSB}
Length composition	year, quarter, stratum, fleet	Log-normal	lnL_{CAL}
Electronic tag (known stock of origin)	stock, year, quarter, stratum, age class	Multinomial	lnL_{ET}
Stock of origin	year, quarter, stratum, movement age class	Normal	lnL_{SOO}

A likelihood weighting scheme (the ω values of equation 8.18, **Table 8.6**) was selected that balanced the contribution of the various data sources, could pass established 'red face tests'.

Table. 8.6. Likelihood weightings for various components of equation 8.18.

Likelihood component	Symbol	Weighting (ω)
Total catches (weight)	ω_c	*0.02*
Index of exploitable biomass (assessment CPUE index)	ωi	*1*
Index of spawning stock biomass (e.g., a larval survey)	ω_{SSB}	*2 (GOM_LAR_SUV & MED_LAR_SUV = 20)*
Length composition	ω_{CAL}	*0.05 or 1* (Reference grid L and H)
Stock of origin	ω_{SOO}	*1*
Electronic tag (known stock of origin)	ω_{ET}	*5*
Recruitment deviations (prior)	ω_{recdev}	*1*
Movement (prior)	ω_{mov}	*1*
Selectivity (prior)	ω_{sel}	*1*
F deviation from master index (prior)	ω_{FD}	*1*
F deviation from master index, annual (prior)	ω_{FA}	*1*
Difference in early/late R0 estimates for recruitment levels 1 and 3.	ω_{Rodiff}	*1*
"Prior" for season trend of biomass in natal area	ω_{natal}	*1*
"Prior" for western stock mixing	ω_{mix}	*1*
"Prior" for scale of mean SSB by area	ω_{SSBmu}	*1*

8.3 Characterising uncertainty

<u>Baseline</u>

Operating models are derived from MLE parameter estimates. Within operating model uncertainty (variability among simulations of the same operating model) occurs only in MSE projections and is based on stochasticity in projected annual recruitment deviations and observation errors

<u>Alternative options</u>

Include within-model uncertainty (parameter uncertainty) via Monte Carlo sampling from the inverse Hessian matrix of model parameters.

Within-model uncertainty via MCMC sampling of posteriors for model parameters.

Concentrate on among-model uncertainty using the maximum posterior density estimates of model parameters and a prior model weight based on expert judgement. Uniform weights will be used to start, possibly updated later using a Delphi-type approach.

9 TRIAL SPECIFICATIONS

9.1 Reference Grid

There are four major uncertainty axes in conditioning and projections in the interim grid: recruitment; natural mortality/maturity (in combination); western stock mixing; scale of the biomass in the East and West areas; weighting of length composition data in the likelihood for operating model conditioning. These axes assume that the options of East and West area (or western and eastern stock) are linked across rows of the table below. This design has the intention of capturing extremes.

Table 9.1. Factors and levels of key uncertainty factors the Reference Grid operating models

Factor: **Recruitment***		
	Western stock	Eastern stock
level **1**	B-H with h=0.6 ("high R0") switches to h = 0.9 ("low R0") starting from 1975	50-87 B-H h=0.98 ("low R0") switches to 88+ B-H *h*=0.98 ("high R0")
level **2**	B-H with h=0.6 fixed, high R0	B-H with h=0.7 fixed, high R0
level **3**	Historically as in level 1. In projections, "low R0" switches back to "high R0" after 10 years	Historically as in level 1. In projections, 88+ B-H with *h*=0.98 ("high R0") switches back to 50-87 B-H with *h*=0.98 ("low R0") after 10 years
Factor: **Spawning fraction/Natural mortality rate for both stocks**		
level **A**	Younger spawning (E+W same)/High natural mortality	
level **B**	Older spawning (different for the 2 stocks)/Low natural mortality (with senescence)	
Factor: **Scale****		
	West area	East area
level **--**	15kt	200kt
level **-+**	15kt	400kt
level **+-**	50kt	200kt
level **++**	50kt	400kt
Factor: **Length composition weighting in likelihood**		
level **L**	0.05	
level **H**	1	

** For recruitment factor level 1 two stock-recruitment relationships are estimated each corresponding to an historical time period. In both the western stock and the eastern stock the steepness of the stock-recruitment curves are specified for these two time periods but unfished recruitment (R0) is re-estimated to capture a regime shift in stock productivity. Recruitment factor level 3 only differs from level 1 in projections, where the estimated regime shift switches back to earlier productivity. Hence, recruitment factor level 3 does not require fitting, historical fit is the same as factor level 1.*

***The scale factor is intended to reflect extremes of area-specific spawning stock biomass based very approximately on the 2017 stock assessment values. The numbers correspond with mean SSB values over the years 1968-2015 in the West area and 1974-2015 in the East areas. The fitting criterion in the conditioning of any OM includes penalty terms to ensure that the output SSB trajectories for the East and West areas for that OM have means over the periods indicated that match the two values applying to that OM as given in the table.*

The western stock recruitment scenarios are intended to capture two alternative hypotheses for historical recruitment. The 'high then low recruitment' hypothesis is captured by level 1, in which a Beverton-Holt stock recruitment relationship with fixed moderate steepness (R0 estimated) shifts to a higher steepness (mimicking the hockey-stick relationship assumed in assessments) after 1975 (second R0 estimated, but lower than for the pre-1975 period). The 'high recruitment' hypothesis is captured by level 2, a Beverton-Holt recruitment relationship with fixed moderate steepness throughout the time series. The third level for western stock recruitment evaluates the robustness of CMPs to a future shift between these alternative recruitment scenarios. In this third scenario recruitment mimics level 1, but 10-years into the projections the higher steepness and lower R0 switches back to moderate steepness and higher R0.

Similarly, the eastern stock recruitment level 1 has two periods of differing unfished recruitment, level 2 assumes a single unfished recruitment value throughout and the third level, as for the western stock, considers a shift between recruitment scenarios after 10 years. Until very recently level 1 (low then high recruitment) was the prevailing hypothesis; however, recent assessments have estimated lower recruitments providing some support for level 2.

The rationale for recruitment level 3 for both stocks is that if recruitment shifts have occurred in the past they could occur in the future as well.

<u>Combinations for Reference Grid</u>

A full cross of (1, 2, 3) x (A, B) x (--, -+, +-, ++) x (L, H), i.e., 48 scenarios in total (only 32 of which require OM fitting since Recruitment levels 1 and 3 differ only in projection years). Discussion will be required regarding whether, in addition to considering results for each of these scenarios individually, they should also be considered for all scenarios in combination, and if so, how the scenarios should be weighted (if at all) in such a combination.

Table 9.2. The factorial design and labelling of the Reference Grid operating models (Recruitment level 3 OMs do not require estimation and change only in projection years)

Length Comp Wt	**L**							
Scale	**--**		**-+**		**+-**		**++**	
Spawn. Frac. / M	**A**	**B**	**A**	**B**	**A**	**B**	**A**	**B**
Recruitment: 1	OM_1	OM_4	OM_7	OM_10	OM_13	OM_16	OM_19	OM_22
Recruitment: 2	OM_2	OM_5	OM_8	OM_11	OM_14	OM_17	OM_20	OM_23
Recruitment: 3	OM_3	OM_6	OM_9	OM_12	OM_15	OM_18	OM_21	OM_24

Length Comp Wt	**H**							
Scale	**--**		**-+**		**+-**		**++**	
Spawn. Frac. / M	**A**	**B**	**A**	**B**	**A**	**B**	**A**	**B**
Recruitment: 1	OM_25	OM_28	OM_31	OM_34	OM_37	OM_40	OM_43	OM_46
Recruitment: 2	OM_26	OM_29	OM_32	OM_35	OM_38	OM_41	OM_44	OM_47
Recruitment: 3	OM_27	OM_30	OM_33	OM_36	OM_39	OM_42	OM_45	OM_48

9.2 Plausibility weighting

The Group agreed to an initial weighting scheme based on the plausibility scores in the poll, with the understanding that once OMs are reconditioned and the Group has reached the finalization tuning stage of the process later in 2022, these weightings could be reconsidered. It was noted that even though the polling indicated that the Group considered an R3 scenario less likely (this scenario includes a shift in recruitment potential after 10 years), they still considered it necessary to capture some type of time-varying productivity to ensure that MPs are robust to such scenarios. Rounding the results of the polling exercise to the nearest 5 percent, the Group converged in the interim on the initial plausibility weights for the OMs as follows:

Table 9.3. Plausibility weights for OMs by factors (rows) and levels (columns)

Factor/Level	**1**	**2**	**3**	**4**
Recruitment	0.40	0.40	0.20	
Spawn/M	0.50	0.50		
Scale	0.30	0.30	0.15	0.25
Length comp	0.50	0.50		

9.3 Robustness trials

All robustness trials were specified as 1 factor changes from four reference set operating models: #1 (1 A -- L), #2 (2 A -- L), #4(1 B -- L), #5 (2 B -- L). An exception was the time varying regime OM which requires recruitment level 3 dynamics, and was applied to reference set OMs: #3 (3 A -- L) and #6 (3 B -- L). The other exception was the intermediate parameter levels for which maturity and survival factor (A/B) is no longer relevant: #1 (1 A -- L), #2 (2 A -- L), #25 (1 A -- H), #26 (2 A -- H).

Table 9.4. Robustness tests, including priority and OMs on which the test is to be conducted.

Priority	Robustness test description	Notes
1	**Western stock growth curve for eastern stock.**	
2	**Catchability Increases**. CPUE-based indices are subject to a 2% annual increase in catchability in the future.	
3	**Unreported overages.** Future catches in both the West and East areas are 20% larger than the TAC as a result of IUU fishing (not known and hence not accounted for by the CMP).	
4	**High western mixing.** The old mixing axis factor level 2: 20% western stock biomass in East area on average from 1965-2016.	Demoted from the reference grid, this provides a yardstick for evaluating whether robustness trials are 'consequential'. Important for setting scale, but not necessarily important for 'does it matter'.
5	**'Brazilian catches'.** Catches in the South Atlantic, including relatively high takes during the 1950s and 60s, are reallocated from the western stock to the eastern stock.	
6	**Time varying mixing**. Eastern stock mixing alternates between 2% and 11% every three years.	
7	**Non-linear indices.** Hyperstability in indices is simulated in projection years.	Imposes a β parameter of 0.5 in projections: $I = qB^{\beta}$
8	**Persistent change in mixing.** Eastern mixing increases from 2% to 11% after 10 projected years.	Was previously a change in western stock mixing before this was shown to be inconsequential to CMP performance. Hence this has been altered to a change in the eastern stock mixing as this will be more likely to be influential.
9	Varying time of regime change in R3.	Two recruitment level 3 scenarios: (3A--L, 3B--L) are subjected to two alternative changes: - a switch back to early recruitment in the East after 5 projection years and 20 projection years for the west; - a switch back to early recruitment in the East after 20 projection years and 5 projections years for the west.
10	Intermediate parameter levels for M, growth, maturity, scale, regime shifts.	The mean of existing high and low scenarios.
11	Zero eastern stock mixing. No Eastern stock in the West area.	Zero eastern mixing, will require substantial further discussion regarding the interpretation. Apply only to the projections.
12	Upweight US_RR_66_144	Upweight US_RR_66_144 until appreciable changes are seen in in OM. Currently NOT available.

Other Robustness trials

1) Probabilistic movement changes
2) Step-changes in catchability
3) Split Med Larval index

"Second round" issues

The following aspects of uncertainty are suggested to be postponed at this time for consideration rather in a "second round":

1) More than two stocks in some OMs
2) Model only a single stock in some OMs
3) Allow for CMPs that set TACs for the whole Atlantic (note that this will require specification of OM components that allocate such catches between West and East areas each year)
4) Use of CAL data in a CMP
5) TACs allocated on a spatially more complex basis than the traditional west and East+Med areas
6) CMP Changes in technical measures affecting selectivity
7) Changes in stock distributions in the future
8) Future changes in proportional allocation of TACs amongst fleets

10 PERFORMANCE MEASURES/STATISTICS

Projections under CMPs will be for 100 years (unless this leads to computational difficulties) commencing in 2023. Prior to that, for projecting for years between the last year of the condition and 2023, the catches will be set equal to the TACs already set, with abundance index data (and, perhaps in a future round, any further monitoring data such as catch-at-length) not yet available for those years being generated as specified under Section 7. Note that considering a period as lengthy as 100 years is not to imply high reliability for projections for such a long time, but to be able take account of transient effects that can persist for some time for a long-lived species.

10.1 Summary measures/statistics

All depletion metrics below are calculated as the spawning stock biomass (SSB) relative to dynamic SSB_0. Dynamic SSB_0 (MacCall et al. 1985) is the spawning biomass that would have occurred if zero catches had been taken historically and in the future, and is therefore impacted by shifts in recruitment expectations. The approach here differs from MacCall et al. (1985) in that it does not account for stochastic recruitment: the dynamic SSB0 is calculated using year-specific estimates of unfished recruitment (depending on the R0 phase in which the model is in each year) assuming that there was zero fishing, i.e., it lags shifts in productivity and does not include annual process error from recruitment. This is so as not to confound its main purpose here, which is to make allowance for transient effects in the dynamics following a recruitment regime shift. Dynamic SSBMSY is calculated using a fixed fraction of SSB_0, taken from the most recent estimates of SSB_{MSY} relative to unfished (i.e., using the steepness parameter assumed for 2018). Since in some operating models R0 is changing over time, the maximum achievable level of stock biomass is also changing and keeping track of dynamic SSB_0 and dynamic SSB_{MSY} provides a realistic yardstick for evaluating management performance.

MSY quantities are calculated for each stock individually (i.e., not a global aggregate MSY) using the standard approach of Botsford (1981) and Walters and Martell (2004) (Box 3.2 of that book) which efficiently calculates equilibrium yields for an age-structured population dynamics model using growth, the stock recruitment relationship and a fishery selectivity at age vector. Since the operating model has multiple fleets model with time varying exploitation rates among fleets, the aggregate mean selectivity (across all fleets, weighted by their catch) in the final year of the historical period (2019) is used in these MSY calculations. Overfishing reference points are calculated on the basis of catch weight divided by total biomass (exploitation rate, U_{MSY}). This does not require the calculation of asymptotic age-selectivity that is challenging for spatial, seasonal, multistock models where temporal variability in recruitment creates transient effects. This U_{MSY} is calculated for each operating model as the fixed exploitation rate that leads to dynamic SSB / SSB_{MSY} = 1 after 50 projection years.

Performance metrics (m), are provided for each simulation i, (n_{sim} = 48, per OM), for each OM (n_{OM} = 48, reference set OMs), stock/area (s) and CMP. These metrics are summarized as either quantiles or means (μ) across all simulations weighted by OM weight (w). For example, for a weighted mean this is calculated:

$$\mu_{s,OM,CMP} = \frac{1}{n_{sim}\sum_{OM=1}^{n_{OM}} w_{OM}} \sum_{OM=1}^{n_{OM}} \sum_{i=1}^{n_{sim}} m_{s,OM,CMP,i} w_{OM} \qquad (10.1)$$

Table 10.1. Performance measures calculated as part of the MSE outputs for each OM and CMP. Performance measures in bold text indicate the key 7 statistics. Statistics are calculated as weighted quantiles and means on a simulation by simulation basis (see Eqn 10.1)

Measure	Measure Description	Statistics*
VarC	Average annual variation in catches (AAV) among CMP update times t (note that except where the resource is heavily depleted so that catches become limited by maximum allowed fishing mortalities, catches will be identical to TACs) defined by: $AAVC = \frac{1}{nt}\sum_{t=1}^{nt} \lvert C_t - C_{t-1}\rvert / C_{t-1}$ (13.1)	Median
AvC10	Mean catches over first 10 projected years. Required to provide short-term vs long-term (AvC30) yield trade-offs.	Median
AvC20	Mean catches over first 20 projected years	Median
AvC30	Mean catches over first 30 projected years	Median
AvgBr	Average Br (spawning biomass relative to dynamic SSB_{MSY}) over projection years 11-30	Median and 5th percentile
Br20	Depletion (spawning biomass relative to dynamic SSB_{MSY}) after projection year 20	Median
Br30	Depletion (spawning biomass relative to dynamic SSB_{MSY}) after projection year 30	Median and 5th percentile
PGT	'Probability Good Trend', 1 minus probability of negative trend (Br31 – Br35) and Br30 is less than 1. Probability of 1 is biologically better. In cases where all simulations are above Br30, PGT = 1 regardless of trend. This allows further discrimination between CMPs that have comparable Br30.	Median
LD	Lowest depletion (spawning biomass relative to dynamic SSB_0) over the 30 years for which the CMP is applied.	5th and 15th percentiles
C1	Catch in first projection year	Median
C20	Mean catches over projected years 11-20	Median
C30	Mean catches over projected years 21-30	Median
D10	Depletion (spawning biomass relative to dynamic SSB_0) after the first 10 projected years	Median
D20	Depletion (spawning biomass relative to dynamic SSB_0) after projection year 20	Median
D30	Depletion (spawning biomass relative to dynamic SSB_0) after projection year 30	Median
DNC	D30 using the MP relative to D30 had no catches been taken over the 30 projected years	Median
LDNC	LD using the MP relative to LD had no catches been taken over the 30 projected years.	Median
POS	Probability of Over-Fished status ($SSB < SSB_{MSY}$) after 30 projected years.	Mean
PNOS	Probability not Over-fished status (1 – POS)	Mean
POF	Probability of Overfishing ($U > U_{MSY}$) after 30 projected years	Mean
PNOF	Probability of not Overfishing (1- POF)	
PGK	Probability of Green Kobe ($SSB > SSB_{MSY}$ & $U < U_{MSY}$) after 30 projected years	Mean
PNRK	Probability of not Red Kobe ($SSB > SSB_{MSY}$ or $U < U_{MSY}$) after 30 projected years	Mean
OFT	'Overfished Trend': Average trend (in log space) of SSB over projection years 31 - 35 when Br30 < 1. $OFT = \begin{cases} 0.1 & SSB_{30} \geq dynSSB_{MSY} \\ m(\log SSB_{31:35}) & SSB_{30} < dynSSB_{MSY} \end{cases}$ Where $m(\vec{x})$ is the gradient of a line of best fit through the vector $\vec{x}$, found via a least squares	Median

** For each of these distributions, 5%-, 50%- and 95%iles are to be reported from 200 replicates. The choice of these percentiles may need further exploration with stakeholders.*

Further stakeholder orientated measures may need to be included. These must be scientifically based, easily understood by stakeholders and such that managers may readily request the evaluation of any changes in options.

10.2 Summary plots

Catch and spawning biomass trajectories plotted as:

a) Annual medians with 5%- and 95%-ile envelopes
b) 10 worm plots of individual realisations

Note that repetitions for different options for selectivity may be needed.

A set of five summary performance metrics S_i, per stock/area are included in a primary results table, these include PGK, AvC10, AvC30, VarC and LD (see Table 10.1 above). To calculate the ranking of CMPs each metric is scaled minimum to maximum and the position in that range is a ranking score R, from 0-1.

For metrics *S,* where small values are better (i.e. VarC) this is calculated:

$$R_{CMP} = \frac{S_{CMP} - min(S)}{max(S) - min(S)} \quad (10.2)$$

For metrics S, where large values are better (e.g. AvC30) this is calculated:

$$R_{CMP} = \frac{max(S) - S_{CMP}}{max(S) - min(S)} \quad (10.3)$$

For example, if CMPs 1-3 obtained VarC values of [1%, 5% and 10%] (lower is better) the ranking scores R would be [0, 0.444, 1]. If CMPs 1-3 obtained AvC30 values of [2, 3, 4] (higher is better) the ranking scores R would be [1, 0.5, 0].

The total rank score *T*, by CMP across metrics *i* (n_i = 5: PGK, AvC10, AvC30, VarC, LD), is the weighted mean of these *R* values by CMP multiplied by a weighting scheme *W* (Table 10.2):

$$T_{CMP} = \frac{1}{\sum_{i=1}^{n_i} W_i} \sum_{i=1}^{n_i} R_{i,CMP} W_i \quad (10.4)$$

Table 10.2. Performance metrics included in a primary results summary table and the various options for weighting of those metrics used to provide an overall 'Total' score.

Examples of weighting schemes	*Status* PGK (mean)	*Yield* AvC10 (50%)	*Yield* AvC30 (50%)	*Stability* VarC (50%)	*Safety* LD* (%TBD)
Default: Equal across yield, stability, and safety	0	0.5	0.5	1	1
Sensitivity 1: Double weighting of safety	0	0.25	0.25	0.5	1
Sensitivity 2: Double weighting of yield	0	1	1	1	1

PGK: Probability of Green Kobe ($SSB > SSB_{MSY}$ & $U < U_{MSY}$) after 30 projected years
AvC10: Mean catches over first 10 projected years
AvC30: Mean catches over first 30 projected years
VarC: Average variation (%) in catches between management periods

10.3 Level of reporting

Baseline

a) Catch-related measures/statistics by traditional West and East+Med areas.
b) Spawning biomass depletions measures/statistics by separate stocks

Alternative options

Many can be conceived, likely related primarily to catch and depletion by some combination of stock and/or spatial stratum. However, these might be left for a "second round", as they would become more pertinent in the face of greater model complexities possibly introduced at that time, such as changing spatial distributions of stocks and/or catches (resulting from changed proportional allocations to different fleets).

A possible additional metric could be VarC but for downward adjustments only.

It may be necessary to characterize stock trajectory to further differentiate among CMPs with similar statistics relating to biomass.

11 REFERENCES

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

Alvarez-Berastegui, D., Tugores, M.P., Ottmann, D., Martín-Quetglas, M., and Reglero, P. 2020. Bluefin tuna larval indices in the western Mediterranean, ecological and analytical sources of uncertainty. SCRS/2020/067. Col. Vol. Sci. Pap. ICCAT. 77(2): 289-311.

Anon. 2014. Report of the 2013 bluefin meeting on biological parameters review (Tenerife, Spain – May 7 to 13, 2013). Col. Vol. Sci. Pap. ICCAT. 70(1): 1-159.

Anon. 2017. Report of the 2017 ICCAT bluefin stock assessment meeting (Madrid, 20-28 July, 2017). Col. Vol. Sci. Pap. ICCAT. 74(6): 2372-2535.

Anon. 2019. Report of the intersessional meeting of Panel 2 (Madrid, Spain, 4-7 March 2019).

Anon. 2021. Report of the first 2021 intersessional meeting of the bluefin tuna Species Group (including W-BFT data preparatory) (Online, 5-13 April 2021).

Arrizabalaga, H., Arregui, I., Medina, A., Rodríguez-Ezpeleta, N., Fromentin, J. M., and Fraile, I. 2019. Life history and migrations of Mediterranean bluefin tuna. The future of bluefin tunas: Ecology, fisheries management, and conservation, 67-93.

Botsford, L. 1981. Optimal fishery policy for size-specific density-dependent population models. J. Math. Biol. 12, 265–293

Bull, B., Francis, R.I.C.C, Dunn, A., McKensie, A., Gilbert, D.J., Smith, M.H., Bain, R., Fu, D. 2012. CASAL (C++ algorithmic stock assessment library): CASAL user manual. V2.30-2012/03/21. NIWA Technical Report. 135. 280pp.

Carruthers, T.R. 2018. Calculating Population-wide spatial and seasonal relative abundance indices for Atlantic bluefin tuna for use in operational modelling. SCRS/2017/019. Col. Vol. Sci. Pap. ICCAT. 74(6): 2586-2595.

Carruthers, T.R. and Butterworth, D.S. 2018. A mixture model interpretation of stock of origin data for Atlantic bluefin tuna. SCRS/2018/133. Col. Vol. Sci. Pap. ICCAT. 75(6): 1363-1372.

Carruthers, T.R., McAllister, M.K., Taylor, N. 2011. Spatial surplus production modelling of Atlantic billfish and tunas. Ecological Applications. 21(7): 2734-2755.

Cort, J. L., Arregui, I., Estruch, V., and Deguara, S. 2014. Validation of the growth equation applicable to the eastern Atlantic bluefin tuna, Thunnus thynnus (L.), using Lmax, tag-recapture and first dorsal spine analysis. Reviews in Fisheries Science & Aquaculture, 22: 3, 239–255.

Fournier, D.A., Hampton, J., Sibert, J.R. 1998. MULTIFAN-CL: a length-based, age-structured model for fisheries stock assessment, with application to South Pacific albacore, Thunnus alalunga. Can. J. Fish. Aqua. Sci. 55: 2105-2116.

Kimoto, A., Carruthers, T., Mayor, C., Palma, C., and Ortiz, M. 2019. Summary of input data (catch and size) used in the Atlantic bluefin tuna Operating Models in 2021. SCRS/2021/031.

Lauretta, M., Walter, J., and Ingram, W. 2018. United States Bluefin tuna index updates: Rod and reel Gulf of Mexico Longline Larval index. SCRS/P/2018/055.

MacCall, A. D., Klingbeil, R. A., and Methot, R. D. 1985. Recent increased abundance and potential productivity of Pacific mackerel. CalCOFI Report, 26: 119–129.

Martell, S. 2015. The iSCAM project. Available at: https://code.google.com/p/iscam-project/

Methot, R.D. and Wetzel, C.R. 2013. Stock synthesis: a biological and statistical framework for fish stock assessment and fishery management. Fish. Res. 142: 86-99.

Quinn, T.J., Deriso, R.B. 1999. Quantitative fish dynamics. Oxford University Press, New York.

Rodriguez-Marin E, Ortiz M, Ortiz de Urbina J.M., Quelle P, Walter J, Abid N, Addis P., Alot E., Andrushchenko I., Deguara S., Di Natale A., Gatt M., Golet W., Karakulak S., Kimoto A., Macias D., Saber S., Santos M. N., and Zarrad R.. 2015 Atlantic bluefin tuna (Thunnus thynnus) Biometrics and Condition. PLoS ONE 10(10): e0141478.doi:10.1371/journal.pone.0141478

Rouyer, T., Bal G., Derridj O., and Fromentin, J.-M. 2021. Update of the French aerial abundance index for 2020 and first attempt at accounting for the environmental effects on bluefin tuna availability in the Gulf of Lions. SCRS/2021/020.

Taylor, N.G., McAllister, M.K., Lawson, G. L. Carruthers, T.R., Block, B.A. 2011. Atlantic bluefin tuna: a novel multistock spatial model for assessment of population biomass. PLoS ONE. 6(12):1:10.

Vázquez Bonales J.A. , Cañadas A., Alemany F., Tensek S., and Pagá García A. 2018. ICCAT GBYP aerial survey for bluefin tuna spawning aggregations in 2018. SCRS/2018/175. Col. Vol. Sci. Pap. ICCAT. 75(6): 1525-1544.

Walters, C.J., and Martell, S.J.D. 2004. Fisheries ecology and management. Princeton University Press.

12 Appendix 1 – year indexing in the OM fitting and ABTMSE R packages

Table App.1.1. The year indexing for the M3 model fitting and R package

Calendar year	ABTMSE indexing (dset)	Historical year	Conditioning year	Projected Year	MP implementation year
1864		1			
1865		2			
1866		3			
...		...			
1963		100			
1964		101			
1965	1		1		
1966	2		2		
...	...		...		
2008	44		44		
2009	45		45		
2010	46		46		
2011	47		47		
2012	48		48		
2013	49		49		
2014	50		50		
2015	51		51		
2016	52		52		
2017	53		53		
2018	54		54		
2019	55		55		
2020	56			1	
2021	57			2	
2022	58			3	
2023	59			4	1
2024	60			5	2
...	...			...	...
2070	106			51	48
2071	107			52	49
2072	108			53	50