2019 Alternative Stock Assessments Report

Patagonian toothfish (*Dissostichus eleginoides*)

Skeljo F
Winter A

Fisheries Department
Directorate of Natural Resources
Falkland Islands Government
Stanley, Falkland Islands

31 May 2020

SA – 2019 – TOO

For citation purposes this publication should be referenced as follows:

Skeljo F, Winter A. 2020. 2019 Alternative Stock Assessments for Patagonian toothfish (*Dissostichus eleginoides*). Fisheries Report SA-2019-TOO. Fisheries Department, Directorate of Natural Resources, Falkland Islands Government, Stanley, Falkland Islands. 22 p.

Distribution: Public Domain

Reviewed and approved by:

Andrea Clausen   Date: 17 June 2020
Director of Natural Resources

Table of Contents

Summary

1. Two data-poor stock assessment models are applied to available data and compared with the established stock assessment for Patagonian toothfish (*Dissostichus eleginoides*) using CASAL software. The two data-poor models are CMSY+ (Froese et al. 2017) and JABBA (Winker et al. 2018); both are Bayesian surplus production models with simple data requirements, and substantially different formulation than the CASAL model.

2. Overall, both alternative models produced similar MSY estimates, moderately different relative biomass estimates, and substantially different absolute biomass estimates compared to the CASAL model. The CMSY+ model proved of limited use due to its sensitivity to assumptions about current stock status. The JABBA model performed better as it makes no assumptions about the current stock status, and its relative biomass and MSY estimates seemed less susceptible to changes in priors.

3. Compared to both alternative models, we consider CASAL as the preferred approach to the toothfish stock assessment in Falkland Islands, as it is an integrated, age-structured model capable of accounting for dynamics in gear selectivity and lagged effects of recruitment and mortality.

4. Findings of this report might provide insight into the usefulness of CMSY+ and JABBA data-poor approaches for other species /stocks managed by FIFD, where data limitations prevent the use of integrated CASAL assessment.

1. Introduction

Patagonian toothfish (*Dissostichus eleginoides*) is a large notothenioid fish found on the southern shelves and slopes of South America and around the sub-Antarctic islands of the Southern Ocean. It is a long-lived species (>50 years), which initially grows rapidly on the shallow shelf areas, before undertaking an ontogenetic migration into deeper waters (Collins et al. 2010). In Falkland Islands waters, Patagonian toothfish spawn on the slopes of Burdwood Bank at ca. 1000 m depth with a minor abundance peak in May, and a major peak in July to August (Laptikhovsky et al. 2006). The eggs, larvae, and small juveniles (<10 cm TL) develop and grow in epipelagic layers of the Falkland Current, and when juveniles attain 10-12 cm TL (<1 year old; Lee 2017), they start to migrate towards the Patagonian shelf and are found at depths <100 m (Arkhipkin and Laptikhovsky 2010). Immature toothfish remain there for 3-4 years, and then, on reaching 60-70 cm TL, migrate into deeper water over the Patagonian slope (Laptikhovsky et al. 2008).

The Falkland Islands exploratory toothfish longline fishery began in 1992 and became an established fishery in 1994 (Laptikhovsky and Brickle 2005). Fishing was traditionally conducted using the Spanish system of longlining (although in the beginning a few vessels used the *Mustad Autoline* system), until the 'umbrella' system was introduced in 2007. The latter system was developed to reduce the loss of hooked toothfish to depredation by cetaceans, with hooks set in clusters and an 'umbrella' of buoyant netting set above each cluster. The umbrella floats above the hooks whilst the gear is on the seabed, but when the gear is recovered, it folds over the hooks and around the fish that has been caught, protecting it from depredation (Brown et al. 2010). Following initial trials in 2007, since 2008 the 'umbrella' system has been adopted by all vessels operating in the fishery in the Falkland Islands.

Although longlining is the only fishery targeting toothfish in Falkland Islands waters, notable quantities are taken as a bycatch in finfish and calamari trawl fisheries. In finfish fishery toothfish is a commercially valuable bycatch, while in calamari fishery it is usually discarded, due to the small size of the specimens (20-30 cm TL). These fisheries exploit different parts of the toothfish population in different areas: longlining occurs on the slope and in deep water, finfish trawling on the shelf

primarily north and west of the Falkland Islands, and calamari trawling on the shelf south and east of the Falkland Islands (Figure 1).

Management of the Falkland Islands toothfish fishery has relied on an integrated age-structured stock assessment implemented in CASAL software (Bull et al. 2012) since 2012. Although CASAL has become the standard approach to toothfish stock assessment in the Southern Ocean (Hillary et al. 2006, Dunn and Hanchet 2010, Ziegler and Welsford 2015), examining alternative models can provide a useful comparison (Cadrin and Dickey-Collas 2015), and either increase our confidence in CASAL outputs, or prompt re-evaluation of its assumptions. This report presents the results of two data-poor toothfish stock assessment approaches, using data up to and including year 2019. The two models used are CMSY+ (Froese et al. 2017) and JABBA (Winker et al. 2018); both are Bayesian surplus production models with simple data requirements, using only a small subset of input data needed for CASAL assessment. The alternative model outputs are compared with the results of integrated age-structured toothfish stock assessment done in CASAL for year 2019 (Skeljo and Winter 2020).
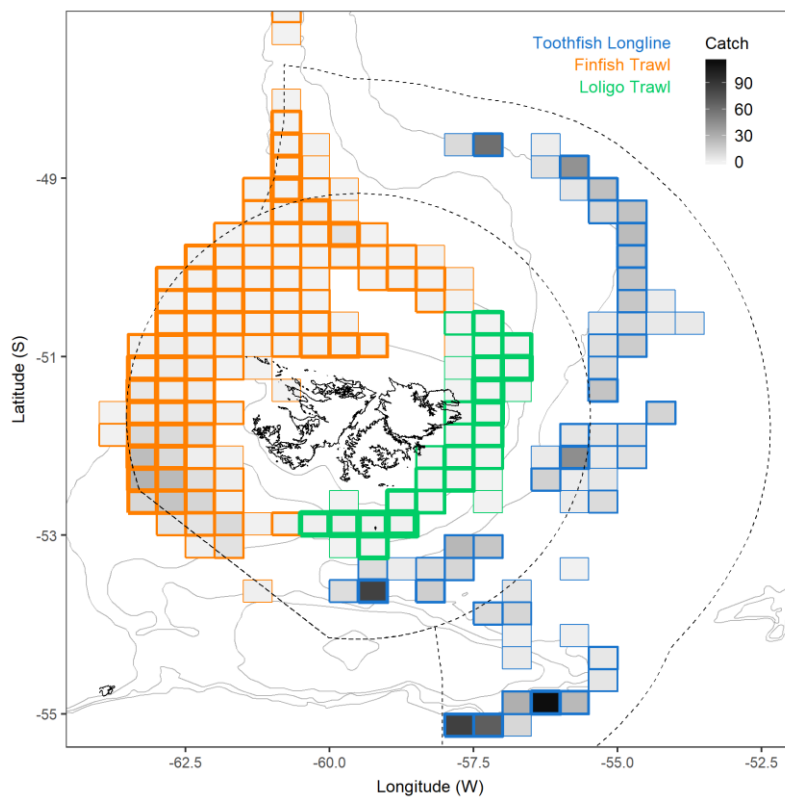

Figure 1. Spatial distribution of toothfish catch and effort by fisheries in 2019. Thickness of grid lines is proportional to the number of vessel days; greyscale is proportional to the toothfish catch biomass (tonnes).

1.1. Stock structure and assumptions

The stock structure of Patagonian toothfish in the Southwest Atlantic is still poorly understood. On a larger spatial scale, there is a well-documented genetic differentiation between toothfish found on the Patagonian Shelf and around South Georgia and South Sandwich Islands (Shaw et al. 2004, Rogers et al. 2006, Canales-Aguirre et al. 2018). However, toothfish population structure across the Patagonian Shelf is less certain, and it is not yet clear whether there are several separate self-sustaining populations or one large meta-population (Parker 2015).

In order to get a better understanding of the toothfish stock structure within Patagonian Shelf (and especially Falkland Islands) waters a range of methodologies were employed by FIFD, most notably: otolith shape analysis, otolith microchemistry analysis and analysis of toothfish

migrations using conventional and satellite tags (Farrugia 2018). Shape analysis revealed significant differences in sagittal otoliths shape between different regions on the Patagonian Shelf, but further evidence is required to identify to what extent this reflects localised stock delineation (Lee et al. 2018). Otoliths microchemistry analysis suggested that juveniles settling on the Falkland Shelf originate from two spatially distinct areas, presumably eastern Burdwood Banks in Falkland Islands waters, and south of Diego Ramirez Islands in Chilean waters (Ashford et al. 2012, Randhawa et al. 2020 *in review*). The existence of separate spawning populations off southern Chile and south of the Falkland Islands on the Burdwood Bank has already been proposed by several authors (Laptikhovsky et al. 2006, Arana 2009, Ashford et al. 2012). Finally, the tagging work done in Falkland Islands waters showed a very high level of site fidelity and limited movement of adult toothfish (Brown et al. 2013), leading to the conclusion that the part of the stock targeted by the longline fishery (primarily older, adult individuals) is most likely confined to Falkland Islands waters.

Considering the currently available information, for the purpose of this assessment we assumed that there is one discrete toothfish stock present in Falkland Islands waters. However, the uncertainty of this assumption must be acknowledged, and should be periodically reviewed to reflect the best available information.

2. Methods

2.1. CMSY+ model setup

CMSY+ is a Bayesian surplus production model, based on the Schaefer production function (Schaefer 1954). It can estimate fisheries reference points, relative stock biomass and exploitation from catch time series and broad priors for the intrinsic rate of population increase r, the carrying capacity K, and the relative stock biomass B/K at the beginning, an intermediate year and the end of the time series.

Since the intrinsic rate of population increase also represents an integrated measure of population resilience, the prior range for r was selected based on the classification of species' resilience in FishBase (Froese and Pauly 2019, Froese et al. 2017). The prior range for K is calculated automatically by default model rules, based on the ratio of highest catch in the time series over r (Froese et al. 2017). A major improvement of CMSY+ over its predecessor CMSY is the introduction of multivariate lognormal priors for r and K, replacing the previous uniform prior distributions and thus simplifying the determination of the 'best' r-K pair (Froese et al. 2019). The prior estimates of relative biomass (B/K) at the beginning and the end of the time series were chosen manually from the set of predefined ranges (Froese et al. 2019), depending on the assumed stock status. The beginning relative biomass range was chosen based on the same assumption used in the integrated CASAL assessment (i.e. that the population was unexploited at the time), and the ending relative biomass range was set to medium level in the base-case scenario, with different assumed ending ranges tested in the sensitivity analysis. The intermediate relative biomass range was calculated automatically by default model rules which are based on the overall catch trend and interplay between the minimum and maximum catch (Froese et al. 2017). Prior ranges used in CMSY+ assessment are given in Table 1.

Once the priors are defined, CMSY+ runs a Monte Carlo algorithm to find the 'viable' r-K pairs. First, a random r-K pair is selected from within the prior ranges for r and K. Then, a starting biomass is selected from the prior biomass range for the first year and the Schaefer surplus production model is used to calculate the predicted biomass in subsequent years. An r-K pair is considered viable if the predicted biomass does not become negative (i.e. stock doesn't crash) and does not fall outside the prior intermediate or ending depletion range. The most probable values of r and K (with 95% confidence limits) are computationally derived from the validated r and K pairs (Froese et al. 2017), and used to calculate standard fisheries reference points such that $MSY =$

$rK/4$, $F_{MSY} = r/2$, $B_{MSY} = K/2$ (Ricker 1975, Schaefer 1954). It should be noted that CMSY+ incorporates a linear decline of surplus production to the Schaefer model when biomass falls below $B_{MSY}/2$ (i.e. $K/4$), as recruitment may be compromised at this level (Haddon et al. 2012, Carruthers et al. 2014, Froese et al. 2015).

This assessment uses the most recent version of CMSY+ R-code available at the time of writing, *CMSY_2019_9f*, published online at http://oceanrep.geomar.de/33076/ in December 2019.

Table 1. Parameter prior ranges used in the CMSY+ model, with a brief description of the selection criteria. The *user defined* prior ranges for resilience and relative biomass at the beginning /end of time series were chosen from the set of predefined ranges suggested by Froese et al (2019).

Parameter	Prior range	Range selection criteria	Description
r	0.05 - 0.5*	user defined	*D. eleginoides* is described as a *'low'* resilience species according to FishBase (Froese and Pauly 2019)
K	24.2 - 72.6*	default	Calculated automatically by default rules, based on the ratio of the highest catch over *r*
B_{1987}/K	0.75 - 1	user defined	Stock is assumed to have been *'nearly unexploited'* at this time, based on the very low reported catches
B_{1998}/K	0.5 - 0.9	default	Calculated automatically by default rules, based on the catch time series trend
B_{2019}/K	0.2 - 0.6	user defined	For the base-case scenario the *'medium depletion'* category was chosen, but other ranges were explored as well

* range is converted into a multivariate lognormal prior in the model

2.2. JABBA model setup

JABBA is a Bayesian state-space surplus production model, based on the generalized Pella-Tomlinson surplus production function (Pella and Tomlinson 1969) of the form:

$$SP_t = \frac{r}{m-1}B_t\left(1 - \left(\frac{B_t}{K}\right)^{m-1}\right),$$

where *r* is the intrinsic rate of population increase at time *t*, *K* is the carrying capacity, *B* is stock biomass at time *t*, and *m* is a shape parameter that determines at which *B/K* ratio maximum surplus production is attained. The Pella-Tomlinson function reduces to the Schaefer function if the shape parameter *m*=2, and to the Fox function if *m* approaches 1. JABBA produces assessment outcomes for all three surplus production functions; for brevity only the Pella-Tomlinson results are presented. The outputs from the other surplus production functions were similar and would not lead to different conclusions.

JABBA can estimate fisheries reference points, relative stock biomass and exploitation from catch and abundance indices time series and priors for the intrinsic rate of population increase *r*, the carrying capacity *K*, and the relative biomass *B/K* at the start of the available catch time series. The major difference of JABBA compared to CMSY is inclusion of relative abundance indices provided in the form of standardized CPUE time-series and assumed to be proportional to biomass. It is important to note that the relative abundance indices available to JABBA are the same ones that are used for CASAL, so there is a predisposition that it should successfully parallel CASAL. The prior range for *r* was based on the classification of species' resilience as provided in FishBase (Froese and Pauly 2019, Froese et al. 2017); the prior for relative biomass *B/K* at the start of the available catch time series was chosen based on the same assumption used in the integrated CASAL assessment (i.e. that the population was unexploited at the time); and the prior for *K* was defined to be similar with the one used in CMSY+ and CASAL assessment (it could not be exactly the same, as priors are defined in slightly different ways between these models). Prior ranges used in JABBA assessment are given in Table 2.

Once the priors were defined, JABBA was executed in R environment using version *v1.5beta* (available online at https://github.com/Henning-Winker/JABBAbeta). The Bayesian posterior distributions of all quantities of interest are estimated by means of a Markov Chains Monte Carlo (MCMC) simulation. Two MCMC chains with 30,000 iterations each were used, with a burn-in of 5,000 for each chain and a thinning rate of five iterations. A full description of the JABBA model, including formulation and state-space implementation, prior specification options and diagnostic tools is available in Winker et al. (2018).

Table 2. Parameter priors used in the JABBA model, with a brief description of the selection criteria.

Parameter	Prior	Description
r	log-normal; range = 0.05 - 0.5*	*D. eleginoides* is described as a *'low'* resilience species according to FishBase (Froese and Pauly 2019)
K	log-normal; μ = 60,000, cv = 1	Defined to be similar with the one used in CMSY+ and CASAL assessment
B_{1987}/K	log-normal; μ = 1, cv = 0.1	Stock is assumed to have been *'nearly unexploited'* at this time, based on the very low reported catches

* range is converted into a lognormal prior in the model

2.3. Data

Three datasets were available as information for the CMSY+ and JABBA stock assessment models: total removals by combined fisheries (required by both models), and catch-per-unit-of-effort (CPUE) time series for Spanish- and umbrella-system longline fisheries (required only by JABBA) (Table 3).

Table 3. Datasets used for the CMSY+ and JABBA stock assessment models.

Data	Time series	CMSY+ input data	JABBA input data
CPUE			
Spanish-system longline	1996-2007, 2013		+
umbrella-system longline	2007-2019		+
Removals	1987-2019	+	+

CPUE

Although CPUE data were available for trawl fisheries as well, only longline CPUE was used as a relative abundance index. This is motivated by the inconsistency of the toothfish CPUE in trawl fisheries, where this species is not targeted, and its bycatch may change due to factors other than stock abundance (e.g. fisheries are switching targets or areas). The longline CPUE data were treated separately for Spanish- and umbrella-system longline, according to the documented difference in the toothfish CPUE between these two fishing gears /techniques (Brown et al. 2010).

For the umbrella-system longline, additional data selection had to be performed in order to avoid introducing bias in the CPUE estimates. The most substantial decision, compared to the previous year's assessment, was to use only the CPUE data from Falkland Islands flagged vessels. The reason is that the fishing was predominantly done by a single Falkland Islands vessel since the onset of the umbrella-system (*CFL Gambler*, replaced by *CFL Hunter* in 2017), assisted occasionally by one or two chartered Chilean vessels. None of the chartered vessels fished in Falkland Islands waters in more than two years since 2007, and their CPUE data were inconsistent. Moreover, at least one of these vessels had restrictions imposed on its fishing practice (e.g. limit on the number of fishing days in the 'best' fishing grounds), which were not in place for the Falkland Islands vessel. All of this led to a conclusion that the CPUE would be more representative as an index of abundance if only Falkland Islands vessels data were used. With a similar goal, data from the 'tagging trips' and from the

longline sets at depths <600 m were removed from the analysis. Tagging trips were removed because part of the actual catch was not reported (corresponding to the tagged and released fish), leading to a biased, lower estimates of CPUE. Fishing in shallow waters was excluded because longlining is prohibited at depths <600 m, and the corresponding sets were not regular commercial fishing (most likely they were experimental fishing, with the aim of collecting brood stock for the toothfish rearing facility).

The CPUE data selected for inclusion in the analysis were prepared for modelling in three steps. First, unstandardized CPUE values were calculated for each longline set as the reported toothfish catch in kg per 1,000 hooks. Second, these were multiplied by the whale depredation rates. Estimation of whale depredation is described in more detail in the *Removals* section of the report, but essentially, toothfish catch depredated from the longline before being hauled on board is also accounted for when calculating CPUE. Since this 'true' catch equals reported + depredated catch, resulting CPUE values will on average be higher than the ones calculated solely from the reported catch. Third, CPUE was standardised using a generalised linear model (GLM), providing a time series of CPUE values (with the associated standard errors) which were assumed to be relative abundance indices (Appendix 1, Appendix 2).

Removals

Total removals were calculated by adding three distinct catch components: (a) reported catches in Falkland Islands waters, (b) catches taken by *Illegal, Unreported and Unregulated* (IUU) fishing, and (c) catches lost to undetected whale depredation.

All reported toothfish catches taken in longline (targeted catch) and trawl fisheries (bycatch) were used, going back to 1987.

The IUU fishing is inherently difficult to estimate (Pitcher et al. 2002, Ainsworth and Pitcher 2005), and no reliable information specific to the Falkland Islands waters was found. Therefore, we utilized the data for the *Antarctic region* from *Table 2* in Agnew et al. (2009), which give estimates of IUU fishing as a percentage of reported catch in 1980-2003. The Antarctic region data was chosen because it pertains specifically to toothfish. Since these data don't cover the whole assessed period, the level of IUU fishing in later years was assumed to be 5% of the reported catches, based on the overall trend of IUU decline post-2000 (CCAMLR 2010).

Whale depredations are included in longline catch reports when they are evident as toothfish hauled up damaged or destroyed by bite-marks. However, toothfish taken entirely by whales before hauling are not accounted for in the reports. In order to address this, Winter and Pompert (2016) modelled whale depredation in Falkland Islands waters by comparing the toothfish CPUE with and without whales present on observed longline sets, using a generalised additive model (GAM). This allowed us to extrapolate the toothfish catch lost to whale depredation for all commercial longline sets, based on the fishing month, longline position, fishing depth, number of hooks set and soak-time. As the GAM is probabilistic, some longline sets obtained model-fit depredation rates <1, implying less toothfish catch in the absence of whale depredation. That outcome is obviously artefactual, and to make the estimates more precautionary, depredation rates for individual longline sets were therefore adjusted upwards by dividing them with the 5th percentile of their own distribution; a value of approximately 0.87 for the Spanish- and 0.96 for the umbrella-system longline.

In order to combine the above-mentioned catch components into total removals by fishery, first the IUU catches were added to the reported catches in each year, and then the undetected whale depredation rate was applied. Effectively, this assumes that reported and IUU catches experience the same average rate of whale depredation. Total removals used in the CMSY+ and JABBA assessment model runs are given in Appendix 2.

3. Results

3.1. CMSY+ model estimates

In the base-case analysis toothfish was considered a low resilience species (Froese and Pauly 2019), with a medium depletion level in the final year of the assessment period. The key CMSY+ output parameters estimated under these assumptions are summarised in Table 4. At the beginning of the time series the estimated biomass (B_{1987}) was already below the carrying capacity K, and by the end of the time series (B_{2019}) it declined to 0.438 K. This estimate of ending relative biomass (B_{2019}/K) should be interpreted with caution, as its 95% confidence interval is wide and spans disparate levels of stock status (Figure 2). Comparison of the catch time series with the estimated MSY shows that the catches (mostly from finfish trawling) were below MSY before the onset of the longline fishery in 1994, followed by a period with generally high catches and a few catch peaks well above MSY 95% confidence intervals (Figure 3). Since 2006, catches have been stable and fluctuated around MSY, coinciding with the introduction of the ITQ system to the longline fishery. According to the model, high catches in the longline fishery in 1994-2005 led to a quick decrease in biomass, which has significantly slowed down once the catches have been reduced in 2006 (Figure 4).

Table 4. Key output parameters estimated by CMSY+ Schaefer production model (r, K, B_{1987}) and resulting calculations of year 2019 biomass and MSY.

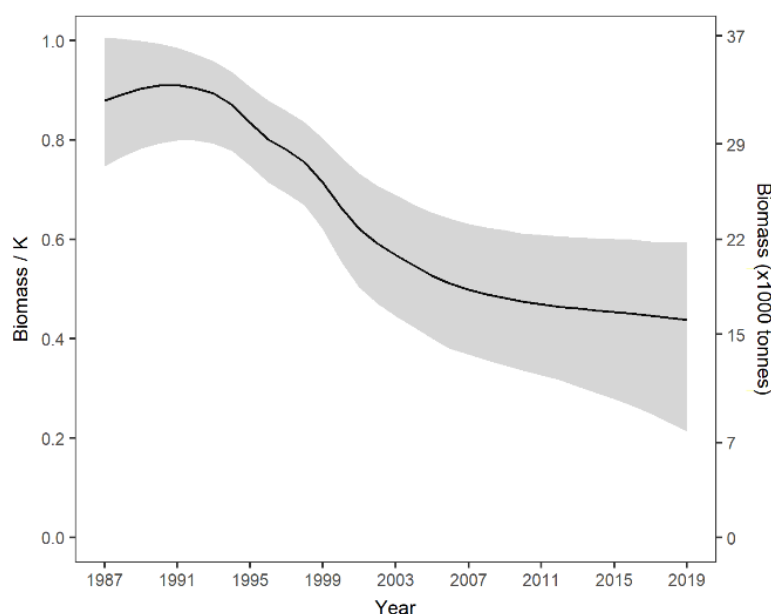| Parameter	median	95% CI
r	0.180	0.098 - 0.329
K	36,667 t	23,967 - 56,097 t
B_{1987}	32,225 t	27,367 - 36,910 t
B_{2019}	16,068 t	7,816 - 21,792 t
B_{1987}/K	0.879	0.746 - 1.007
B_{2019}/K	0.438	0.213 - 0.594
MSY	1,604 t	1,095 - 2,194 t

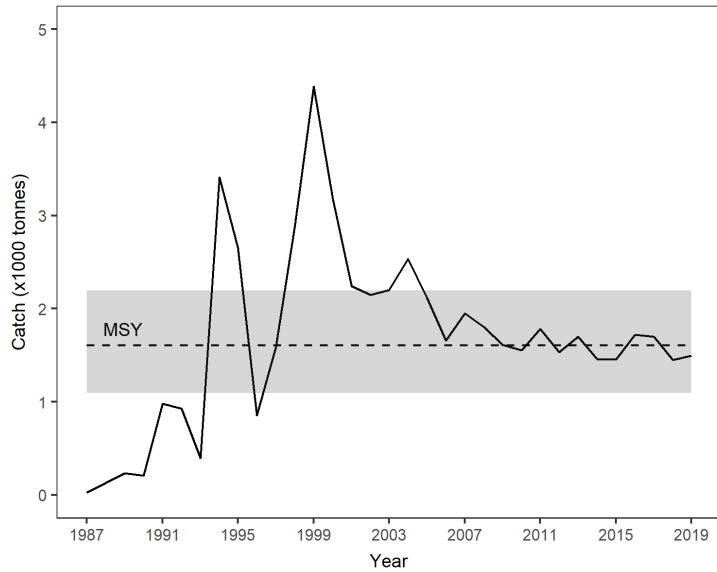Figure 2. Median and 95% confidence intervals of the relative biomass trend estimated by CMSY+.

Figure 3. Toothfish catches (full line), relative to CMSY+ estimated MSY (dashed line) with 95% confidence interval (shaded area).
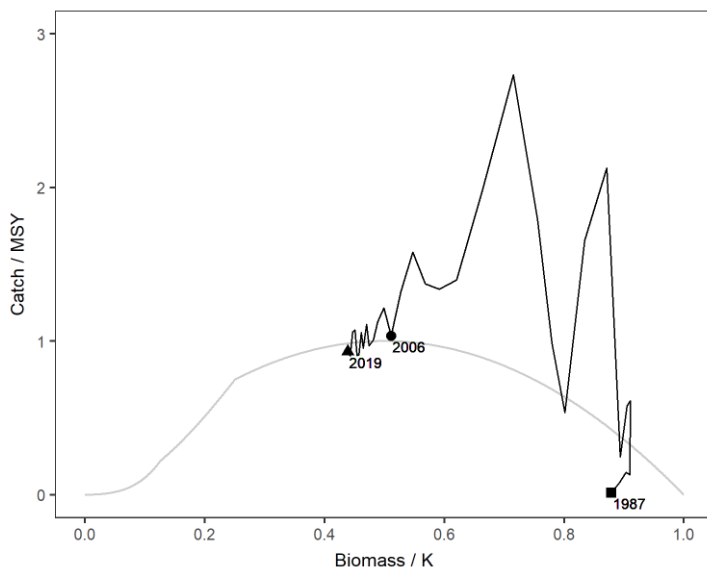

Figure 4. CMSY+ estimates of catch/MSY relative to *B/K*, from beginning (square) to the ending (triangle) year in the time series. The Schaefer equilibrium curve, indented at *B/K* < 0.25 to account for reduced recruitment at low stock sizes, is added for reference (grey line); this curve indicates relative catches that would maintain the biomass (relative catches above the curve will shrink future biomass, and relative catches below the curve allow future biomass to increase). Year 2006 (circle) marks the introduction of the ITQ system to the longline fishery.

Retrospective analysis
The retrospective analysis was done by successively removing one, two and three final years of data from the base-case model and rerunning the analysis. All three runs produced almost identical relative biomass estimates as the base-case model, with negligible departures (Figure 5). This is a desirable pattern, indicating that the model isn't overly influenced by the observations in the most recent years. It was the expected outcome, as catches are the only observation used by the model, and they have fluctuated very little in the last several years.
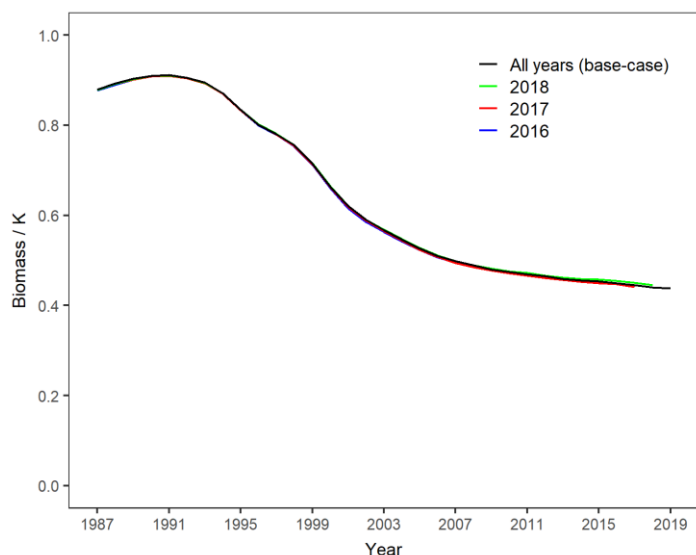
Figure 5. Estimated relative biomass trends for the base-case model (black line) and three retrospective model runs. The numeric label indicates the year up to and including which each individual retrospective model was run.

Sensitivity analysis

One of the requirements of CMSY+ model is defining a prior range for the relative stock biomass at the end of the time series (B_{2019}/K). However, the estimate of the median B_{2019}/K is one of the key model outcomes as well, and is calculated only from the results that fall within the specified prior B_{2019}/K range. This means that the model cannot 'recover' from an incorrectly set ending biomass range, because, by design, ending biomass estimates outside the prior range are discarded by the CMSY+ algorithm (Froese et al. 2017). This leads to a concern that the outcome might be overly influenced by the assumed prior (circular logic), leading to a biased estimate.

In order to explore the sensitivity of our model to the assumed ending biomass range, the base-case scenario was compared to six alternative model runs with different prior ending ranges (Figure 6). It is evident that even a small change of the prior range can lead to a very different median B_{2019}/K estimate, and consequently to a different current stock status estimate. As expected, higher prior ranges lead to a more optimistic assessment outcome, and vice versa.
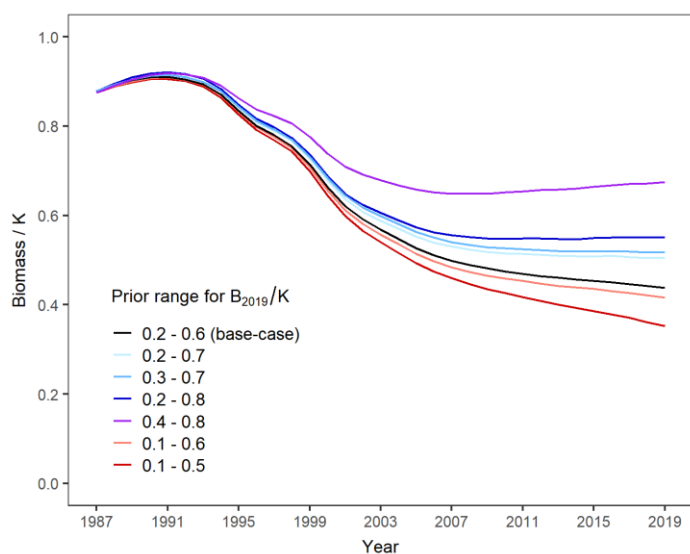

Figure 6. Estimated relative biomass trends for the base-case model (black line) and six alternative model runs with different assumed prior ranges of ending relative biomass (B_{2019}/K).

9

3.2. JABBA model estimates

The key output parameters estimated by JABBA are summarised in Table 5. The estimated biomass declined from 0.953 *K* in 1987 to 0.329 *K* in 2019. Although the ending relative biomass is somewhat low, it should be noted that its 95% confidence interval is wide (0.189 - 0.582), spanning values that would lead to a very different interpretations of stock status (Figure 7). Comparison of the catch time series with the estimated MSY is given in Figure 8, and closely resembles the results of the CMSY+ model. The high catches in 1994-2005 (well above MSY in most years) led to a quick decline in toothfish biomass; however, the introduction of ITQ system to the longline fishery in 2006 lead to a more sustainable catches (fluctuating closely around the median MSY), and the decline in biomass has gradually slowed down, before reverting to an increasing trend in the final year (Figure 9).

Table 5. Key output parameters estimated by JABBA model (r, K, B_{1987}) and resulting calculations of year 2019 biomass and MSY.

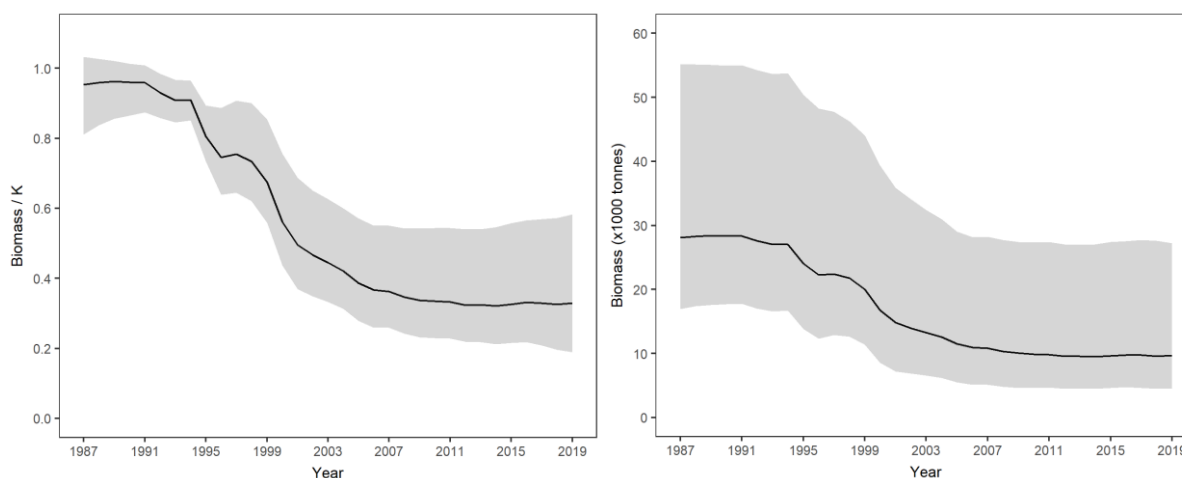| Parameter	median	95% CI
r	0.162	0.067 - 0.294
K	29,753 t	18,614 - 58,031 t
B_{1987}	28,094 t	16,933 - 55,147 t
B_{2019}	9,683 t	4,485 - 27,245 t
B_{1987}/K	0.953	0.811 - 1.032
B_{2019}/K	0.329	0.189 - 0.582
MSY	1,632 t	1,109 - 2,093 t

Figure 7. Median and 95% confidence intervals of the relative (left) and absolute (right) biomass trend estimated by JABBA. Note that the median trends are identical (although shown on a different scale), and only the confidence intervals differ.
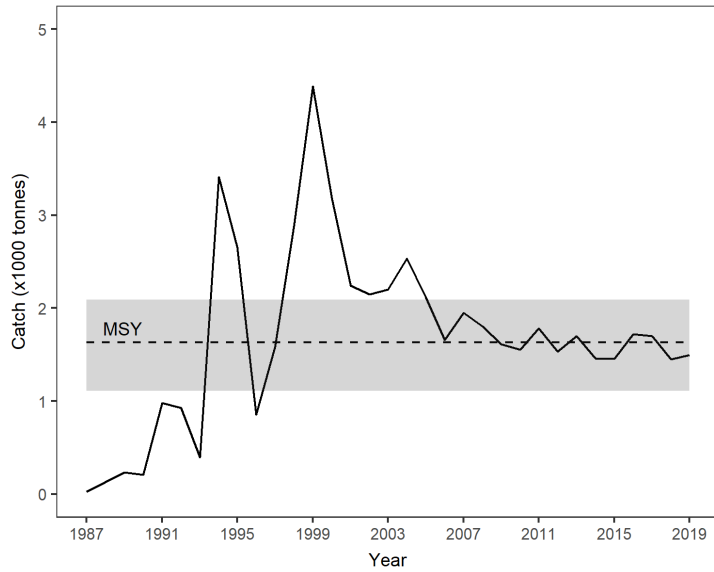
Figure 8. Toothfish catches (full line), relative to JABBA estimated MSY (dashed line) with 95% confidence interval (shaded area).
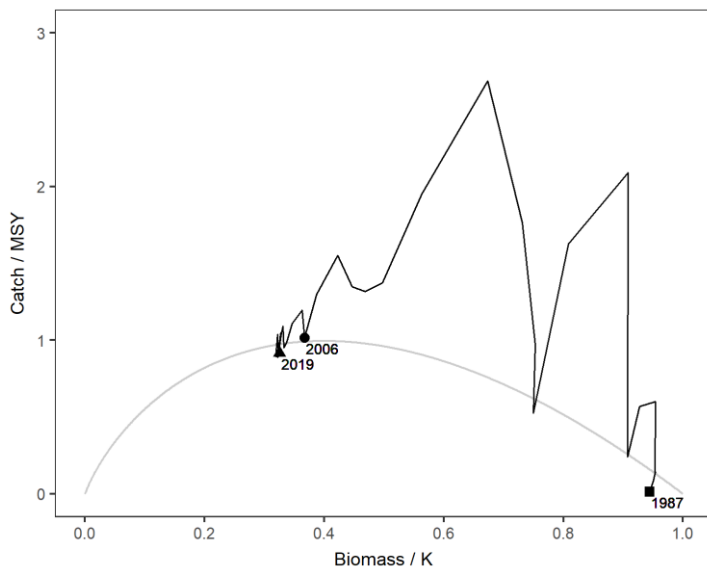
Figure 9. JABBA estimates of catch/MSY relative to *B/K*, from beginning (square) to the ending (triangle) year in the time series. Pella-Tomlinson equilibrium curve, with inflection point at 0.40 *K*, is added for reference (grey line); this curve indicates relative catches that would maintain the biomass (relative catches above the curve will shrink future biomass, and relative catches below the curve allow future biomass to increase). Year 2006 (circle) marks the introduction of the ITQ system to the longline fishery.

The retrospective analysis was done by successively removing one to six final years of data from the base-case model and rerunning the analysis. All six runs produced similar relative biomass estimates as the base-case model, and no systematic trend in departures from the base-case model was evident (Figure 10). Unlike CMSY+, JABBA uses both catch and CPUE observations, which is why there is comparatively more variation in retrospective analysis - CPUE showed higher fluctuations in recent years than did the catches.
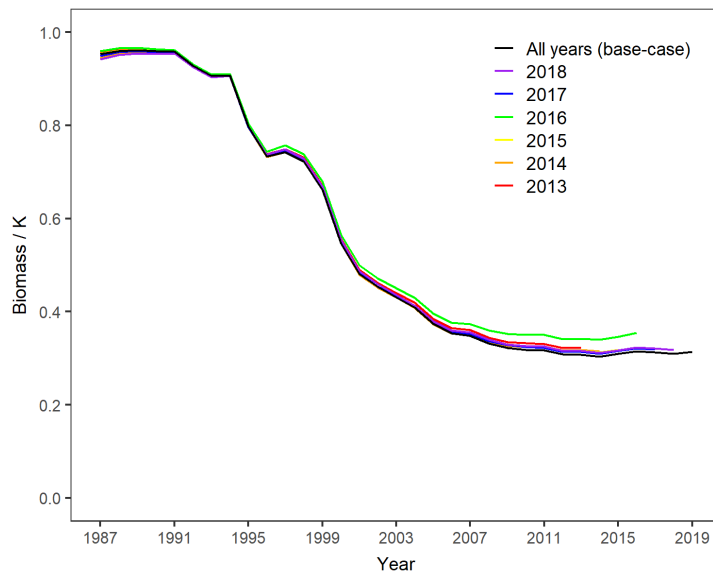

Figure 10. Estimated relative biomass trends for the base-case model (black line) and six retrospective model runs. The numeric label indicates the year up to which individual retrospective model was run (inclusive).

Sensitivity analysis

Priors for *r*, *K* and relative biomass at the beginning of the time series (B_{1987}/K) need to be specified before JABBA model can be run. Here we tested the effect of different prior settings on the model outcomes; although this is by no means an extensive analysis, it does suggest in which way each of the priors affects the model, and can be considered as a first step towards a more comprehensive analysis in the future.

In the first scenario, mean of the lognormal prior for *K* was decreased from 60,000 to 40,000 t; compared to the base-case model, this change had almost no effect on the estimated relative biomass and only a minor effect on the absolute biomass (Figure 11). In the second scenario, *cv* of the lognormal prior for B_{1987}/K was increased from 10% to 25%; this led to a notably lower estimates of relative and absolute biomass at the beginning of the time series, but as the model progressed the trend became increasingly similar to the base-case model, with only a minor difference in the final year. In the third scenario, higher intrinsic population growth rate was assumed, and *r* range was increased from 0.05-0.5 to 0.25-0.57 (both ranges are found on FishBase). This change had the most substantial effect on the model outcomes and produced a distinct trend compared to the other scenarios, as the model now described the stock in terms of higher *r* and much lower *K*. In other words, instead of explaining the observed data by a larger stock size with low productivity (as in the base-case), the model did it by a smaller stock size with high productivity. The effect of the higher prior *r* on the relative biomass estimates was less obvious, producing very similar result to the other scenarios in the final year (although, unlike other scenarios, with an increasing trend in the recent years). The absolute biomass estimates were approximately 40% lower than in the base-case, but high estimated *r* meant that the population could produce high surplus biomass, leading to a higher estimated MSY.

Overall, this sensitivity analysis indicates that the assumptions about species intrinsic population growth rate /resilience have the most impact on the model outcomes and should be further explored. However, base-case and all three alternative scenarios produced comparable relative biomass estimates in the final year, as well as similar MSY.
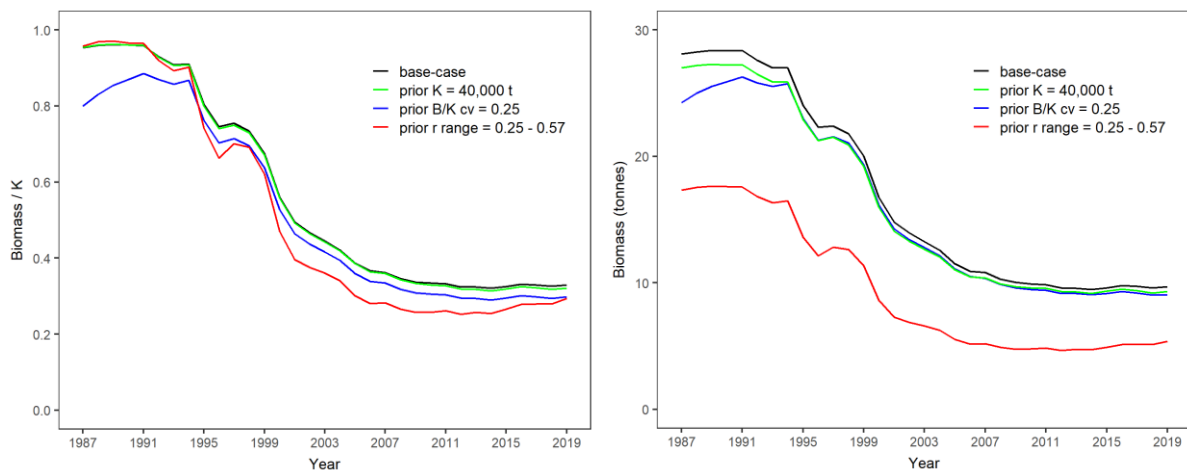

Figure 11. Estimated relative (left) and absolute (right) biomass trends for the base-case model (black line) and three alternative model runs with different assumed priors for K, r and B_{1987}/K.

3.3. Comparison of CMSY+, JABBA and CASAL assessment

Since CMSY+, JABBA and CASAL models have very different structure, and estimate a wide variety of output parameters, not all their outcomes can be directly compared. Assessment in CASAL is the preferred approach to toothfish stock assessment in Falkland Islands, so here we present only the CMSY+ and JABBA model outcomes which are directly comparable with the CASAL results. It is important to point out that CASAL does not provide estimates of K and B_{2019}/K, as the model outcomes are expressed in terms of the spawning stock biomass instead of total biomass. To make the comparison with CMSY+ and JABBA possible, we approximated B_{2019}/K by B_{2019}/B_{1987} for the CASAL model (i.e. we treated the total biomass in 1987 as being at the carrying capacity). This was based on the CASAL estimate of spawning stock biomass in 1987 being only slightly lower than the estimated unexploited spawning stock biomass; extending the same logic to the total biomass didn't seem unreasonable.

Overall, CMSY+ and JABBA produced lower estimates of both relative and absolute biomass throughout the time series compared to CASAL (Table 6, Figure 12). This is especially pronounced for the absolute biomass, with alternative models estimates at approximately half of the CASAL estimate at the beginning of the time series, and half (CMSY+) or one-third (JABBA) at the end of the time series. The trends are different as well, with the alternative models showing a steep decline approximately in 1994-2006 and levelling off afterwards, and CASAL estimate exhibiting a constant moderate decline throughout the time series. In general, differences in relative biomass estimates between CASAL and the alternative models are much lower than the differences in absolute biomass. The MSY estimates are more similar, with alternative models producing approximately 15% lower values than CASAL.

Table 6. Comparison of the selected output parameters estimated by CASAL, CMSY+ and JABBA stock assessment models for the toothfish in Falkland Islands waters (1987-2019). All biomass estimates are given in tonnes.

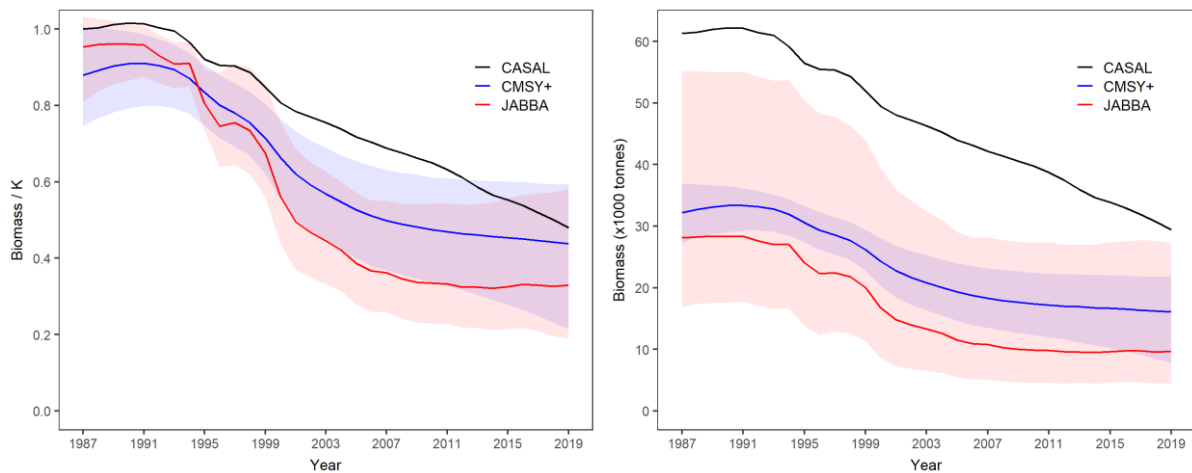|	CASAL	CMSY+	JABBA
K	~ B_{1987}	36,667 (23,967 - 56,097)	29,753 (18,614 - 58,031)
B_{1987}	61,280 (54,298 - 268,511)	32,225 (27,367 - 36,910)	28,094 (16,933 - 55,147)
B_{2019}	29,392 (23,668 - 244,378)	16,068 (7,816 - 21,792)	9,683 (4,485 - 27,245)
B_{2019}/K	0.480*	0.438 (0.213 - 0.594)	0.329 (0.189 - 0.582)
MSY	1,890 (1,665 - 7,205)	1,604 (1,095 - 2,194)	1,632 (1,109 - 2,093)

* calculated as B_{2019}/B_{1987}

Figure 12. Estimated relative (left) and absolute (right) biomass trends for CASAL, CMSY+ and JABBA models.

4. Discussion

The two alternative models used for the toothfish stock assessment in this report, CMSY+ and JABBA, belong to the surplus production models (SPM). SPMs are among the least data demanding population models that can produce estimates of MSY and associated fisheries reference points, and despite a number of limitations (Maunder 2003, Punt and Szuwalski 2012), remain an integral tool for data-limited to -moderate stock assessments (Dichmont et al. 2016, Punt et al. 2015). The main limitations of SPMs is that they ignore the stock's size/age structure and therefore fail to account for dynamics in gear selectivity (Wang et al. 2014) and lagged effects of recruitment and mortality (Aalto et al. 2015, Punt and Szuwalski 2012), which can both lead to biased assessment results. However, SPMs have been considerably enhanced by the introduction of Bayesian methods with improved prior formulations, development of frameworks that allow incorporating both observation and process errors, and Bayesian state-space modelling approaches (Winker et al. 2018).

In our analysis, CMSY+ estimated similar MSY and relative toothfish biomass in the final year as did CASAL, but its absolute biomass estimate was substantially lower. The main shortcoming of CMSY+ assessment was its sensitivity to the changes in prior for the ending relative stock biomass. As already described, design of the model is such that only the estimates of ending relative biomass which fall within its assumed prior range are kept, making it impossible for the model to recover from incorrectly set prior (i.e. incorrect assumption of the current stock status). To an extent this can be mitigated by selecting broad prior ranges (Froese et al. 2017), but our sensitivity analysis showed that even relatively small changes of the prior range can lead to a very different current stock status

estimate, making the usefulness of CMSY+ model for Falkland Islands toothfish stock assessment questionable.

JABBA biomass estimates were lower than those obtained by CMSY+, and even less resembled the CASAL estimates. One of the advantages of JABBA over CMSY+ is that it doesn't require a prior for relative biomass at the end of the time series, avoiding the need for knowledge of the current stock status. It does require a relative biomass prior for the beginning of the time series, but sensitivity analysis indicated that differences in starting relative biomass tend to decrease as the model progresses towards the final year, making the outcomes more robust against the assumed prior. However, the prior for the species intrinsic growth rate r had a large impact on the model outcomes and requires more comprehensive exploration. Besides testing of different priors, JABBA assessment would also benefit from any improvements in the CPUE data; further options of standardizing this data will be explored as a part of the future CASAL stock assessment, and they should be introduced to the JABBA model as well.

Overall, both alternative models produced similar MSY estimates, moderately different relative biomass estimates, and substantially different absolute biomass estimates compared to the CASAL model. CMSY+ proved of limited use due to its sensitivity to the assumptions about current stock status. JABBA makes no assumptions about current stock status, and its relative biomass and MSY estimates seemed less susceptible to changes in priors. The JABBA model has potential for further use, preferably accompanied by an extensive sensitivity analysis. Compared to both alternative models, we consider CASAL as the preferred approach to toothfish stock assessment in Falkland Islands, as it is an integrated, age-structured model capable of accounting for dynamics in gear selectivity and lagged effects of recruitment and mortality. However, it needs to be stressed that the comparison between these models wasn't straightforward, as they have very different formulations. That being said, all three models produced MSY estimates that are above the annual toothfish catches taken in the last two years, adding a measure of confidence that the stock was exploited in a sustainable manner. Lastly, findings of this report might provide insight into the usefulness of CMSY+ and JABBA data-poor approaches for other species/stocks managed by FIFD, where data limitations prevent the use of integrated CASAL assessment.

5. References

Aalto EA, Dick EJ, MacCall AD. 2015. Separating recruitment and mortality time lags for a delay-difference production model. Canadian Journal of Fisheries and Aquatic Sciences 165, 161-165.

Agnew DJ, Pearce J, Pramod G, Peatman T, Watson R, Beddington JR, Pitcher TJ. 2009. Estimating the worldwide extent of illegal fishing. PLoS ONE 4, e4570. https://doi.org/10.1371/journal.pone.0004570

Ainsworth CH, Pitcher TJ. 2005. Estimating illegal, unreported and unregulated catch in British Columbia's marine fisheries. Fisheries Research 75, 40-55.

Arana P. 2009. Reproductive aspects of the Patagonian toothfish (*Dissostichus eleginoides*) off southern Chile. Latin American Journal of Aquatic Research 37, 381-394.

Arkhipkin AI, Laptikhovsky VV. 2010. Convergence in life-history traits in migratory deep-water squid and fish. ICES Journal of Marine Science 67, 1444-1451.

Ashford JR, Fach BA, Arkhipkin AI, Jones CM. 2012. Testing early life connectivity supplying a marine fishery around the Falkland Islands. Fisheries Research 121-122, 144-152.

Brown J, Brickle P, Hearne S, French G. 2010. An experimental investigation of the 'umbrella' and 'Spanish' system of longline fishing for the Patagonian toothfish (*Dissostichus eleginoides*) in the Falkland Islands: Implications for stock assessment and seabird by-catch. Fisheries Research 106, 404-412.

Brown J, Brickle P, Scott BE. 2013. Investigating the movements and behaviour of Patagonian toothfish (*Dissostichus eleginoides* Smitt, 1898) around the Falkland Islands using satellite linked archival tags. Journal of Experimental Marine Biology and Ecology 443, 65-74.

Bull B, Francis RICC, Dunn A, McKenzie A, Gilbert DJ, Smith MH, Bian R, Fu D. 2012. CASAL (C++ algorithmic stock assessment laboratory): CASAL User Manual v2.30-2012/03/21. NIWA Technical Report 135, 275 p.

Cadrin SX, Dickey-Collas M. 2015. Stock assessment methods for sustainable fisheries. ICES Journal of Marine Science 72, 1-6.

Canales-Aguirre CB, Ferrada-Fuentes S, Galleguillos R, Oyarzun FX, Hernández CE. 2018. Population genetic structure of Patagonian toothfish (*Dissostichus eleginoides*) in the Southeast Pacific and Southwest Atlantic Ocean. PeerJ 6, e4173.

Carruthers TR, Punt AE, Walters CJ, MacCall A, McAllister MK, Dick EJ, Cope J. 2014. Evaluating methods for setting catch limits in data-limited fisheries. Fisheries Research 153, 48-68.

CCAMLR. 2010. Estimation of IUU catches of toothfish inside the convention area during the 2009/10 fishing season. Document WG-FSA-10/6 Rev. 1, 12 p.

Collins MA, Brickle P, Brown J, Belchier M. 2010. The Patagonian toothfish: biology, ecology, and fishery. Advances in Marine Biology 58, 227-300.

Dichmont CM, Deng RA, Punt AE, Brodziak J, Chang YJ, Cope JM, Ianelli JN, Legault CM, Methot RD, Porch CE, Prager MH, Shertzer KW. 2016. A review of stock assessment packages in the United States. Fisheries Research 183, 447-460.

Dunn A, Hanchet SM. 2010. Assessment models for Antarctic toothfish (*Dissostichus mawsoni*) in the Ross Sea including data from the 2006–07 season. New Zealand Fisheries Assessment Report 2010/1, 28 p.

Farrugia TJ. 2018. Stock discrimination research overview for Patagonian toothfish in the Falkland Islands. Fisheries Department, Directorate of Natural Resources, Falkland Islands Government, 13 p.

Froese R, Demirel N, Coro G, Kleisner KM, Winker H. 2017. Estimating fisheries reference points from catch and resilience. Fish and Fisheries 18(3), 506-526.

Froese R, Demirel N, Coro G, Winker H. 2019. A simple user guide for CMSY+ and BSM (CMSY_2019_9f.R). Published online at http://oceanrep.geomar.de/33076/ in December 2019.

Froese R, Demirel N, Sampang A. 2015. An overall indicator for the good environmental status of marine waters based on commercially exploited species. Marine Policy 51, 230-237.

Froese R, Pauly D (Editors). 2019. FishBase. World Wide Web electronic publication. www.fishbase.org, version (12/2019), accessed at www.fishbase.org in May 2020.

Haddon M, Klaer N, Smith DC, Dichmont CD, Smith ADM. 2012. Technical Reviews for the Commonwealth Harvest Strategy Policy. FRDC 2012/225. CSIRO. Hobart. 69 p.

Hillary RM, Kirkwood GP, Agnew DJ. 2006. An assessment of toothfish in subarea 48.3 using CASAL. CCAMLR Science 13, 65-95.

Laptikhovsky VV, Arkhipkin AI, Brickle P. 2006. Distribution and reproduction of the Patagonian toothfish *Dissostichus eleginoides* Smitt around the Falkland Islands. Journal of Fish Biology 68, 849-861

Laptikhovsky VV, Arkhipkin AI, Brickle P. 2008. Life history, fishery and stock conservation of the Patagonian toothfish, around the Falkland Islands. American Fisheries Society Symposium 49, 1357-1363.

Laptikhovsky VV, Brickle P. 2005. The Patagonian toothfish fishery in Falkland Islands' waters. Fisheries Research 74, 11-23.

Lee B, Brewin PE, Brickle P, Randhawa H. 2018. Use of otolith shape to inform stock structure in Patagonian toothfish (*Dissostichus eleginoides*) in the south-western Atlantic. Marine and Freshwater Research 69, 1238-1247.

Maunder MN, Punt AE. 2004. Standardizing catch and effort data: A review of recent approaches. Fisheries Research 70, 141-159.

Maunder MN, Starr PJ. 2003. Fitting fisheries models to standardised CPUE abundance indices. Fisheries Research 63, 43-50.

Maunder MN. 2003. Is it time to discard the Schaefer model from the stock assessment scientist's toolbox? Fisheries Research 61, 145-149.

Parker SJ. 2015. Stock discrimination tools for Falkland Islands toothfish. NIWA Client Report NEL2015-002, 28 p.

Pella JJ, Tomlinson PK. 1969. A generalized stock production model. Inter-American Tropical Tuna Commission Bulletin 13, 421-458.

Pitcher TJ, Watson R, Forrest R, Valtýsson HÞ, Guénette S. 2002. Estimating illegal and unreported catches from marine ecosystems: a basis for change. Fish and Fisheries 3, 317-339.

Punt AE, Su N-J, Sun C-L. 2015. Assessing billfish stocks: a review of current methods and some future directions. Fisheries Research 166, 103-118.

Punt AE, Szuwalski C. 2012. How well can FMSY and BMSY be estimated using empirical measures of surplus production? Fisheries Research 134-136, 113–124.

Randhawa HS, Lee B, Brickle P, Reid MR, Arkhipkin AI. 2020. Oceanographic cues determine the recruitment demography of toothfish *Dissostichus eleginoides* on the Patagonian Shelf: evidence from otoliths microchemistry. *In review*.

Ricker WE. 1975. Computation and Interpretation of Biological Statistics of fish Populations. Bulletin of the Fisheries Research Board of Canada 191, Ottawa, Canada, 382 p.

Schaefer M. 1954. Some aspects of the dynamics of populations important to the management of the commercial marine fisheries. Bulletin of the Inter-American Tropical Tuna Commission 1, 27-56.

Shaw PW, Arkhipkin AI, Al-Khairulla H. 2004. Genetic structuring of Patagonian toothfish populations in the southwest Atlantic Ocean: the effect of the Antarctic polar front and deep-water troughs as barriers to genetic exchange. Molecular Ecology 13, 3293-3303.

Skeljo F, Winter A. 2020. 2019 Stock assessment report for Patagonian toothfish (*Dissostichus eleginoides*). Fisheries Report SA-2019-TOO. Fisheries Department, Directorate of Natural Resources, Falkland Islands Government, Stanley, Falkland Islands. 36 p.

Wang S-P, Maunder MN, Aires-da-Silva A. 2014. Selectivity's distortion of the production function and its influence on management advice from surplus production models. Fisheries Research 158, 181-193.

Winker H, Carvalho F, Kapur M. 2018. JABBA: Just another Bayesian biomass assessment. Fisheries Research 204, 275-288.

Winter A, Pompert J. 2016. Initial analysis of whale depredation in the Falkland Islands toothfish longline fishery. Fisheries Department, Directorate of Natural Resources, Falkland Islands Government, 18 p.

Ziegler P, Welsford D. 2015. An integrated stock assessment for the Heard Island and the McDonald Islands Patagonian toothfish (*Dissostichus eleginoides*) fishery in Division 58.5.2. Document WG-FSA-15/52. CCAMLR.

Appendix 1. CPUE standardization

Spanish- and umbrella-system longline CPUE was standardized using generalized linear model (GLM), with a log link function and normally distributed error (Maunder and Starr 2003, Maunder and Punt 2004). Individual longline haul CPUE values (expressed as toothfish catch in kg per 1000 hooks) were the response variable, and the explanatory variables considered in the model are given in Table A.1.

Table A.1. Explanatory variables considered in the CPUE standardization GLM, by fishery and type.

Explanatory variables		Variable type
Spanish-system	umbrella-system	
Year*	Year*	Categorical
Month*	Month*	Categorical
Region*	Region*	Categorical
Depth	Depth	Continuous
Soak-time*	Soak-time*	Continuous
Vessel*	-	Categorical
-	Hooks-per- umbrella*	Categorical
-	Umbrella-spacing	Categorical

* Variables which were found statistically significant and included in the final model.

The *Month* variable accounts for the seasonal variability in CPUE, and the *Region* variable attempts to capture the spatial distribution of CPUE, divided into three broad areas: (a) within the Falklands zone and south of 53.5° S (Burdwood Bank spawning area), (b) within the Falklands zone and north of 53.5° S, and (c) outside the Falklands zone. *Depth* variable is the average fishing depth at which longline is set (in meters). *Soak-time* was calculated in hours-per-hook for Spanish-system longline, and hours-per-line for the umbrella-system. *Vessel* variable was excluded from the umbrella-system longline CPUE standardization, as the only two vessels used in the assessment never fished concurrently in the same year, making the *Vessel* and *Year* effects indistinguishable. The umbrella-system had two additional variables: *Umbrella-spacing* (which was changed from 40 m between umbrellas to 22 m between umbrellas after November 2014) and number of *Hooks-per-umbrella* (which was progressively decreased from 10 hooks initially to 8 hooks in December 2007, to 7 hooks in March 2014, to 6 hooks in June 2016).

Year effect is the quantity of interest so it must be a part of the final CPUE model, and the remaining explanatory variables were added to the *Year* by forward stepwise selection, and included in the final model only if they improved R^2 by at least 0.5%.

Fitting GLM to the Spanish-system data showed that the explanatory variables *Year*, *Month*, *Region*, *Soak-time* and *Vessel* are statistically significant, although the model explained only 17.1% of the overall variation in CPUE. Standardized and unstandardized CPUE time series showed similar trends, with high values in the first 4-5 years of fishery, followed by the lower, but relatively steady values in the later years (Figure A.1).

Fitting GLM to umbrella-system data showed that the explanatory variables *Year*, *Month*, *Region*, *Soak-time* and *Hooks-per-umbrella* are statistically significant, and the model explained 28.5% of the overall variation in CPUE. Comparison of the umbrella-system standardized and unstandardized annual CPUE indexes is shown in Figure A.2. The most prominent feature of the unstandardized data is steep increase in CPUE in 2017, followed by the decline during the next two years, but still with significantly higher values than in the earlier years of fishery. This corresponds to the entry of the new vessel into the fishery, i.e. *CFL Hunter* replaced the *CFL Gambler* from the beginning of 2017 (as mentioned before, only the data belonging to these two vessels were used in the analysis). Second trend is less obvious as it is partially masked by the mentioned 'new vessel' feature, but broadly speaking, there was an increase in unstandardized CPUE from 2014 to 2019. This can be explained by the decrease in the number of hooks-per-umbrella, introduced voluntarily

by vessel operators over time. Since hooks are set in clusters, reducing their number from 8 to 7 to 6 didn't affect the catches per umbrella much, but it was perceived as the reduced effort (calculated as the total number of hooks per longline set) and lead to an increase in unstandardized CPUE. However, the number of hooks-per-umbrella was significant explanatory variable in GLM, and in the standardized CPUE time series this increasing trend was removed. It is worth pointing out that the option of using the umbrellas instead of hooks as the unit of effort was explored as well, but the results were almost exactly the same as when using hooks and having hooks-per-umbrella as a significant explanatory variable in GLM.

The distribution of the residuals from the GLM fit to Spanish- and umbrella-system data was consistent with the assumption of normality (Figure A.3).
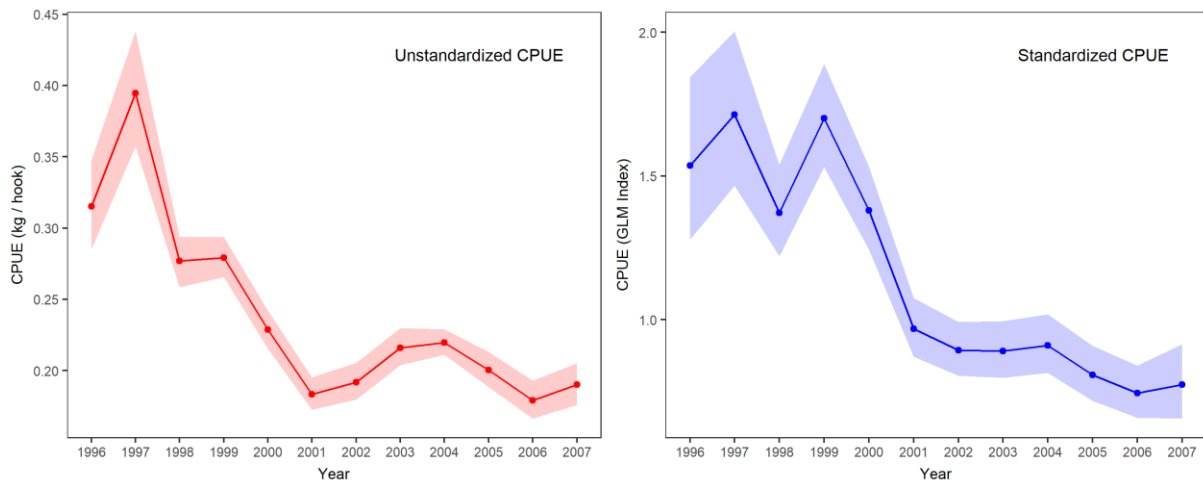

Figure A.1. Spanish-system longline CPUE time series: unstandardized CPUE expressed as toothfish catch in kg per hook (left), and standardized CPUE indices from the GLM (right); shaded areas correspond to 95% confidence intervals.
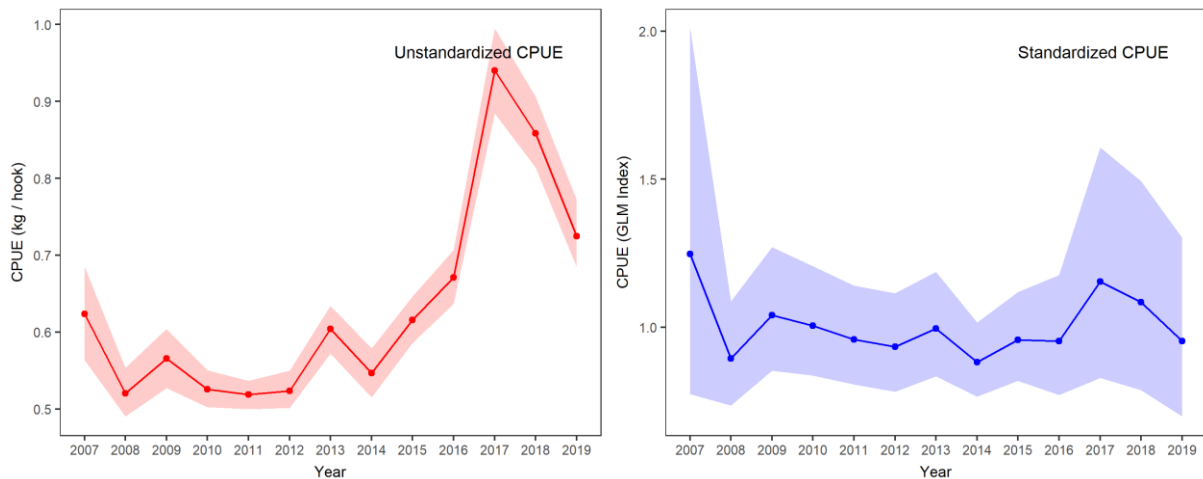

Figure A.2. Umbrella-system longline CPUE time series: unstandardized CPUE expressed as toothfish catch in kg per hook (left), and standardized CPUE indices from the GLM (right); shaded areas correspond to 95% confidence intervals.
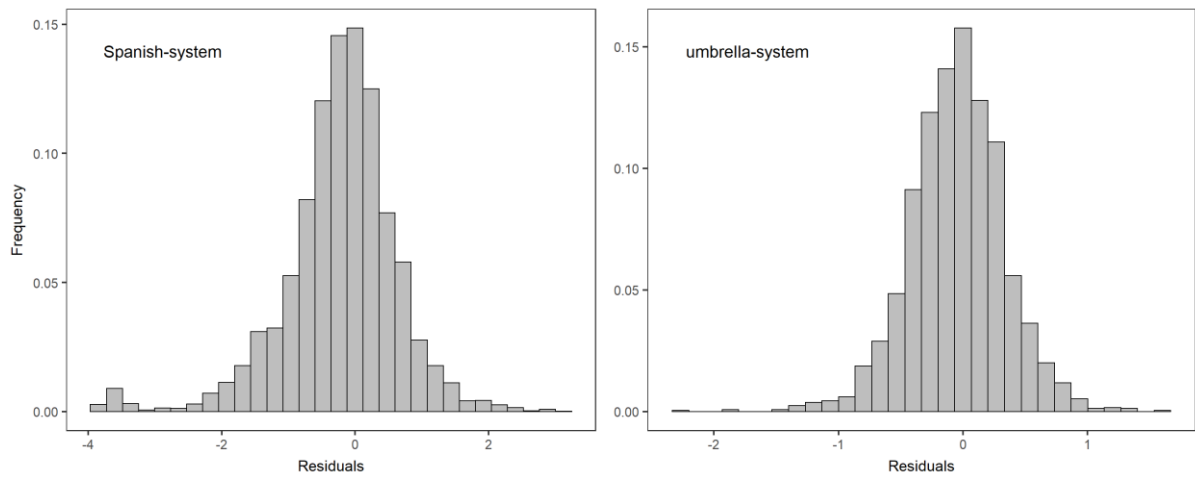
Figure A.3. Density histograms of residuals from the generalized linear model (GLM) fitted to the Spanish- and umbrella-system longline CPUE data.

Appendix 2. Input parameters

Table A.2. Total toothfish removals for combined fisheries, and standardized CPUE indices with standard errors for Spanish- and umbrella-system longline fisheries.

Year	Removals (tonnes)	CPUE indices		CPUE standard errors	
		Spanish-system	umbrella-system	Spanish-system	umbrella-system
1987	23	-	-	-	-
1988	127	-	-	-	-
1989	235	-	-	-	-
1990	208	-	-	-	-
1991	980	-	-	-	-
1992	926	-	-	-	-
1993	394	-	-	-	-
1994	3411	-	-	-	-
1995	2656	-	-	-	-
1996	856	1.536	-	0.092	-
1997	1585	1.714	-	0.078	-
1998	2882	1.371	-	0.058	-
1999	4385	1.700	-	0.053	-
2000	3185	1.380	-	0.052	-
2001	2242	0.968	-	0.053	-
2002	2147	0.893	-	0.052	-
2003	2201	0.891	-	0.055	-
2004	2533	0.911	-	0.056	-
2005	2119	0.808	-	0.059	-
2006	1659	0.744	-	0.061	-
2007	1950	0.774	1.247	0.083	0.238
2008	1806	-	0.895	-	0.097
2009	1613	-	1.041	-	0.100
2010	1552	-	1.005	-	0.091
2011	1782	-	0.959	-	0.087
2012	1530	-	0.934	-	0.088
2013	1697	0.361	0.995	0.388	0.088
2014	1462	-	0.882	-	0.071
2015	1456	-	0.957	-	0.078
2016	1718	-	0.953	-	0.106
2017	1702	-	1.154	-	0.165
2018	1449	-	1.084	-	0.160
2019	1495	-	0.954	-	0.156