Fisheries Research 243 (2021) 106098

Contents lists available at ScienceDirect

Fisheries Research

journal homepage: www.elsevier.com/locate/fishres

Are tag-based integrated stock assessments robust to IUU fishing?

Brett Stacy [a,b,*], Paul Burch [a], Philippe E. Ziegler [c], Katherine A. Cresswell [a,b], Klaas Hartmann [b], Richard M. Hillary [a]

[a] CSIRO Marine and Atmospheric Research, GPO Box 1538, Hobart, Tasmania, 7001, Australia
[b] University of Tasmania, Institute for Marine and Antarctic Studies, Private Bag 49, Hobart, Tasmania, 7001, Australia
[c] Department of the Environment, Australian Antarctic Division, 203 Channel Highway, Kingston, Tasmania, 7050, Australia

ARTICLE INFO

Handled by Steven Xavier Cadrin

Keywords:
Integrated stock assessment
IUU fishing
Tag-recapture
Patagonian toothfish
Catch bias

ABSTRACT

Integrated stock assessments depend on reliable catch and observational data to produce unbiased estimates of fish stock biomass and productivity. However, biases in catch data due to unreported catch are common among many of the world's fisheries, especially those that are high value and therefore more vulnerable to illegal fishing. Some of these fisheries rely on tag-recapture data to support the estimation of biomass in integrated assessments and set precautionary catch limits intended to safeguard the fish stocks from overexploitation. Tag-recapture data is generally considered to be a more powerful indicator of stock abundance than survey data if the underlying assumptions are met. The effect of unreported catches that include tagged fish on biomass estimates in tag-based integrated assessments is unclear. We used a simulation analysis to determine the impact of under-reported catch on a hypothetical Patagonian toothfish (*Dissostichus eleginoides*) stock. Several catch scenarios are presented that cover different magnitudes and trend types of under-reported catch. We used the CASAL model (commonly employed to assess toothfish stocks) as both the operating model and the estimation model to evaluate how the various scenarios affected estimation performance of biomass depletion. Biomass depletion was increasingly underestimated as the magnitude of under-reporting increased, regardless of trend in catch. When unreported catch exceeds twice the amount reported combined with unreported tag returns, the estimation model will fail to detect biomass has been depleted below a target reference point of 50 % within 20 years of fishing. The lack of detection will falsely indicate current catch limits are sustainable when in reality they are not. Importantly, these estimates remained largely unchanged among different trends in under-reported catch. For assessments reliant on tag-recapture data, this suggests that it is more critical to estimate the overall quantity of unreported catch and the number of unreported tagged fish than the trend.

1. Introduction

Contemporary stock assessment models integrate population and fishery information including biological data, catch history, and indices of abundance to produce estimates of life history parameters, stock biomass, and mortality rates (Fournier and Archibald, 1982; Hilborn and Walters, 1992; Maunder and Piner, 2014; Maunder and Punt, 2013; Polacheck et al., 2006). Biomass estimates are often used to evaluate the performance of the fishery and set sustainable catch limits. However, mis-specified parameters and observations can lead to biases in the assessment and estimated sustainable catch limits, potentially leading to either overfishing or underutilization of the resource (Hordyk et al., 2019; Ichinokawa et al., 2014; Wang et al., 2009).

The catch history over an assessment period, in combination with an index of stock abundance, provides crucial information required by most assessment models to estimate stock biomass and productivity (Magnusson and Hilborn, 2007). An incorrect catch history can misrepresent the true fishing mortality the stock has experienced. For surplus production models that include a survey index of abundance, this has been shown to lead to biased estimates of quantities such as spawning stock biomass (SSB) that are commonly used to set sustainable catch limits (Rudd and Branch, 2017; Van Beveren et al., 2017).

Catch data that are used in an assessment can be biased due to many reasons. Removals due to recreational (Zeller et al., 2008) or artisanal (Van der Elst et al., 2005) fishing are often difficult to estimate and hence account for in fishery stock assessments. Discarding fish due to minimum size limits or market demands is a common practice and can result in substantial fishing mortality as discard rates are often high

* Corresponding author at: CSIRO Marine and Atmospheric Research, GPO Box 1538, Hobart, Tasmania, 7001, Australia.
E-mail address: brett.stacy@utas.edu.au (B. Stacy).

https://doi.org/10.1016/j.fishres.2021.106098
Received 18 May 2021; Received in revised form 28 July 2021; Accepted 29 July 2021
Available online 3 August 2021
0165-7836/

(Borges et al., 2005; Hammond and Trenkel, 2005; Roda et al., 2019) and discard survival low (Evans et al., 1994). If left unaccounted for in the reported catch, discarded fish can lead to biased assessment outcomes (Punt et al., 2006). Fishing induced depredation of fish by marine predators can cause significant mortality (Clark and Agnew, 2010; Tixier et al., 2020, 2010), resulting in an underestimate of total removals caused indirectly from fishing activity. This discrepancy can also lead to underestimates of catch rates if not accounted for and hence bias assessments that use commercial catch rates as an index of abundance (Roche et al., 2007). For tagged based assessments, this would have implications for the number of tagged fish that are perceived to be available for recapture (Welsford and Ziegler, 2013).

Catch from IUU (Illegal, Unreported and Unregulated) fishing has previously been estimated to be over ten million tonnes annually (Agnew et al., 2009). IUU catch can include removals due to legal practices (some listed previously) or illegal practices. In the context of this study, IUU catch encompasses both tagged and untagged fish that are removed from the stock and not reported and therefore not available to be included in an integrated assessment. IUU catch can be substantial, but due to its cryptic nature it is extremely difficult to estimate and can therefore pose unique challenges to managers of individual stocks. For example, overcatch of southern bluefin tuna between 1985–2005 was estimated at approximately 178,000 t, discrediting historical estimates of standardized catch rates and size distributions of the catch (Polacheck, 2012). As these quantities are essential elements of the stock assessment, the indication of status and subsequent management decisions made during this period were brought into question. There are many other examples worldwide where IUU catch was not initially identified and included in assessments, leading to biased estimates of stock status (Cisneros-Montemayor et al., 2013; Pitcher et al., 2002; Plagányi et al., 2011; Varkey et al., 2010). Moroccan catches were underestimated for demersal and pelagic species by as much as 50 % during the last three decades of the 20th century (Pitcher et al., 2002). A study estimating under-reported catch of reef fish, tuna, anchovy, shark, sea cucumber, and lobster in Eastern Indonesia found that IUU catch exceeded reported catch for all species by more than a factor of 1.5 (Varkey et al., 2010). IUU catch of abalone in South Africa in 2008 was estimated to exceed the total allowable catch (TAC) by more than 10 times, representing an extreme case (Plagányi et al., 2011). In Mexico the lack of catch logbook validation for legal fleets combined with illegal fishing led to official statistics under-reporting catch by a factor of two from 1950 to 2010 (Cisneros-Montemayor et al., 2013). Fisheries with evidence of IUU catch often require extensive catch reconstruction efforts to develop more useful indicators of stock status (Pauly and Zeller, 2016). In some instances, catch time series are so mistrusted that catch data is left out of an assessment entirely and fishery independent surveys are relied on instead (Cook, 2013; Porch et al., 2006). While there are cases of over-reported catch (Watson and Pauly, 2001), in this study we focus on under-reporting since this appears to be more common and possibly consequential among world fisheries (Pauly and Zeller, 2016).

When not accounted for, or accounted for inaccurately, IUU catch can lead to an overly optimistic assessment of fishery sustainability (Cisneros-Montemayor et al., 2013; Polacheck, 2012). Previous studies have examined the impact of misreported catch on population parameter estimates. Rudd and Branch (2016) simulated several misreported catch scenarios (constant under-reporting, constant over-reporting, increasing reporting rate, and decreasing reporting rate) to determine their effect on estimated population parameters and fishery management metrics using the Pella-Tomlinson surplus production model. They combined survey index data simulated from an operating model with misreported catch to investigate the impact on biomass estimates. There was inconsequential bias in the estimation of SSB depletion when constant misreporting was present, leading to no misestimation of fishing impact under these conditions. However, a gradual change in misreporting rates over time could lead to either underutilization (when reporting rates improve) or overfishing (when reporting rates degrade). Omori et al. (2016) investigated how misreported catch impacts fishery sustainability when it is combined with biased, fisheries-dependent, index of relative abundance (catch rates) in a surplus production model. They examined several case species and came to similar conclusions about constant misreporting leading to unbiased estimates of depletion. Van Beveren et al. (2020) used management strategy evaluation to test the performance of different harvest control rules under past and future catch under-reporting. An integrated, age-structured model using an unbiased total egg production index was applied to Western Atlantic mackerel to demonstrate that rebuilding the stock from a low biomass state was unlikely under any of the harvest control rules they evaluated. Rebuilding would only be possible if future unreported catch decreased to unrealistic low levels and estimates of uncertainty in unreported catch improved substantially. In contrast to Rudd and Branch (2017) and Omori et al. (2016); Van Beveren et al. (2020) only investigated decreasing to constant trends in unreported catch and did not focus on the impact of model mis-specification of catch or unbiased indices of abundance on estimates of depletion.

The impact of under-reported catch on tag-based integrated assessments may be different than integrated assessments that do not use tag-recapture data or non-integrated models like the surplus production approach. Integrated analysis (tag-based or otherwise) incorporate age- or length-structured information such as catch composition (e.g. catch-at-age data) and are considered the preferred method for stock assessments (as opposed to surplus production models for example) (Punt et al., 2020). Temporal changes in composition data collected only from vessels who report their catch, combined with the reported catch, may mislead an assessment if there is significant unreported fishing occurring at the same time (Beare et al., 2005; Cook, 2013). Age- or length-structured models that integrate composition and other information can detect these types of discrepancies through model fitting to data, while surplus production models cannot. Additionally, tag-based assessments have the advantage (compared to the more common catch rate or survey-based assessments) of an absolute, rather than relative, index of abundance through the tag-recapture data (Pine et al., 2003). Unlike relative indices, tag-recapture methods do not rely on estimating a catchability coefficient to inform absolute abundance. Catchability can be influenced by factors such as changes in water temperature, management, abundance, and physiological processes making it difficult to estimate (Patterson et al., 1993; Wilberg et al., 2009; Ziegler et al., 2003). Given these advantages, tag-based integrated assessments may be better suited to reflect changes in abundance than commonly employed alternatives.

There are simplified models that use only tag-recapture data to estimate fish abundance that do not use the integrated approach such as the Lincoln-Peterson or Schnabel models (Ricker, 1975). These models rely on basic information such as number of fish tagged, and number of fish recaptured both with and without tags to estimate abundance. The assumptions for these models are that (1) the population is closed to additions (recruitment or immigration) or deletions (deaths or emigration), (2) recapture probability is equal among all fish in each sample, and (3) tagged fish are not lost or overlooked (Pine et al., 2003). Considering these assumptions in the context of IUU catch, if the proportion of tagged to untagged fish removed illegally is similar to the proportion removed legally, a Lincoln-Petersen estimate of biomass is unlikely to be biased as a result of illegal fishing alone (Welsford and Ziegler, 2013). However, it remains unclear whether tag-based integrated assessments can provide unbiased estimates of biomass when supplied with under-reported catch along with under-reported tag-recaptures. Investigating this influence using age-structured integrated models in general is a logical next step to surplus production models (Rudd and Branch, 2017). Incorporating tag-recapture data as an indicator of abundance is also a logical choice as it has the potential to compensate for illegal fishing without requiring a scaling coefficient to be estimated. The authors have been unable to find evidence in the peer reviewed literature that the influence of unreported catch combined with the unreported removal of tagged fish on parameter estimates and derived quantities from an integrated assessment model has been critically examined.

We examine the effects of under-reported catch on the example of Antarctic and sub-Antarctic toothfish (*Dissostichus* spp.) fisheries where tag-based integrated stock assessment models have been used and IUU fishing is known to have occurred (Agnew, 2000). Patagonian toothfish (*D. eleginoides*) and Antarctic toothfish (*D. mawsoni*) stocks are managed by the Commission for the Conservation of Antarctic Marine Living Resources (CCAMLR). CCAMLR has an extensive fish tagging program where fish are tagged and released during normal fishing operations (Mormede and Dunn, 2013; Ziegler, 2013). Data from tag-releases and subsequent recaptures of toothfish, together with age and or length composition information, estimates of biological parameters, and the history of catch removals, are then integrated in tag-based stock assessment models to estimate stock biomass and set total allowable catch limits using the CCAMLR decision rule (Hanchet et al., 2015). The CCAMLR decision rule relates the current estimate of spawning stock biomass ($SSB_{current}$) to the estimated virgin spawning stock biomass (SSB_0), and catch limits are set such that the ratio of $SSB_{current}$ to SSB_0 (SSB_{status}) will be at 50 % at the end of a 35 year projection period, and has less than a 10 % probability of dropping below 20 %. (Constable et al., 2000).

Toothfish are high value fish with a life expectancy of over 50 years (Collins et al., 2010; Farmer et al., 2019), qualifying them as the type of fish that are typically vulnerable to overfishing from illegal fishing (Plagányi et al., 2011). There has been extensive IUU fishing for toothfish in many fisheries within the CCAMLR area (CCAMLR, 2018a), where in some years, estimated IUU catch was two (Fig. 1.b,c,d) and even three (Fig. 1.a) times greater than reported catch. While efforts have been made to reconstruct catch removals from IUU fishing to include in assessments (Agnew and Kirkwood, 2005), uncertainty in these catch estimates remains high. IUU fishing for toothfish is believed to have effectively ceased since 2011, and CCAMLR has since stopped estimating it (CCAMLR, 2018b, 2018c, 2018a, 2018d). It is unlikely that the true amount of IUU catch will ever be accurately quantified and therefore the construction of a catch history will continue to pose problems in future stock assessments.

This study uses an age-structured, integrated, stock assessment model that relies on tag-recapture data to estimate SSB_0 (virgin, pre-fishing spawning stock biomass), $SSB_{current}$ (spawning stock biomass at the time of the assessment), and SSB_{status} ($SSB_{current}$/ SSB_0) to investigate the estimation performance when provided with biased catch information. Characteristics of estimated SSB were chosen as a performance metric for the model because they are the primary indicator of stock health used in most established fisheries. We use parameters and data from a Patagonian toothfish-like fishery as a case study, comparing simulated SSB depletion of a "true" population of fish (i.e., the operating model) with that estimated by an assessment model (i.e., the estimation model) with prescribed levels of under-reported catch. We also compare the integrated results to those using a Lincoln-Petersen estimator to provide an indication for the cause of the biases we found.

2. Methods

2.1. modelling overview

We used the C++ Algorithmic Stock Assessment Laboratory (CASAL) program (Bull et al., 2012) to investigate the effects of under-reported catch and removal of tagged fish on estimated SSB for a hypothetical toothfish stock. CASAL is routinely used to assess toothfish stocks in the Southern Ocean and many New Zealand fisheries. In addition to its

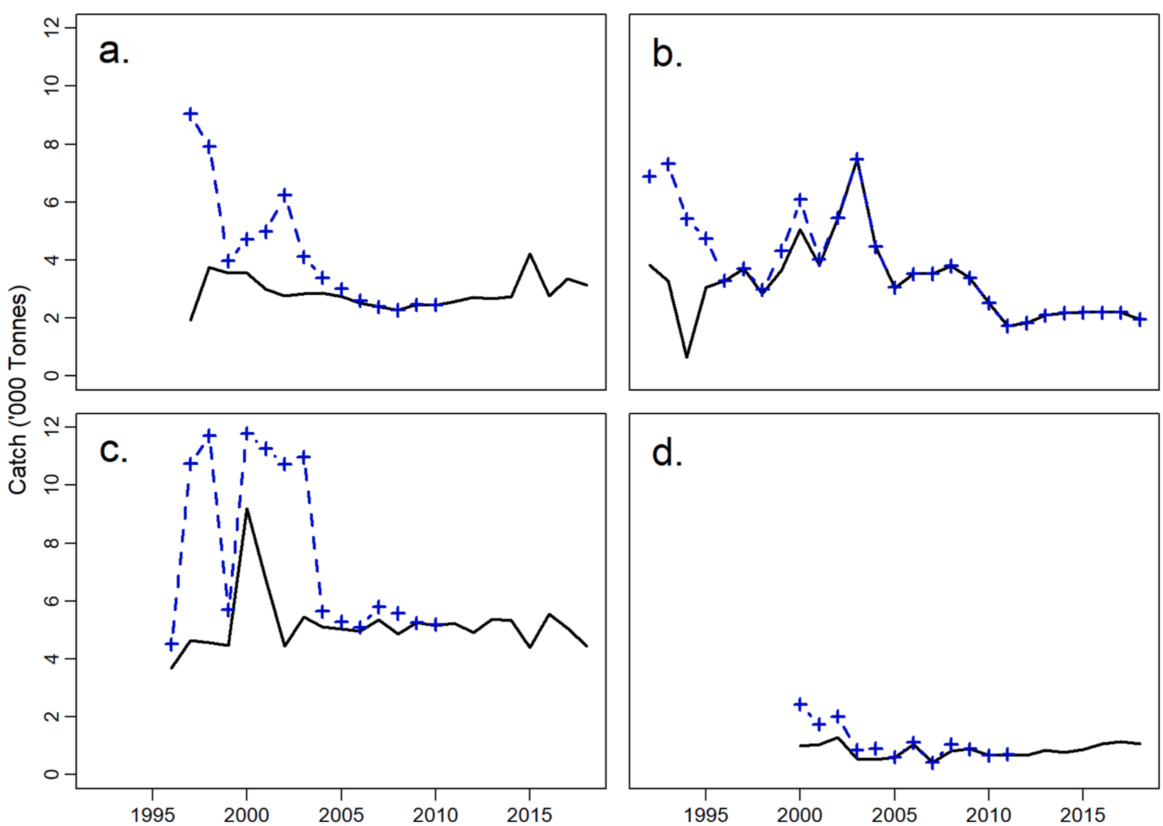

Fig. 1. Catch histories for four Patagonian toothfish fisheries in the sub-Antarctic: a. Heard Island and McDonald Islands (CCAMLR, 2018a), b. South Georgia (CCAMLR, 2018b), c. Kerguelen Islands (CCAMLR, 2018d), and d. Crozet Island (CCAMLR, 2018c). Black lines indicate reported catch (legal only) and blue dashed lines with crosses represent estimated reconstructed catch (legal and IUU catch). Reconstructed catch is not shown for years during which IUU fishing was not estimated.

primary function of acting as an estimation model, CASAL has the capability to simulate data that are consistent with parameters defined by the user. We leveraged this capability to use CASAL as both an operating model and estimation model to evaluate the potential impacts of IUU fishing on a tag-based assessment (Fig. 2).

The operating model was used to generate a "true" time series of SSB for each catch scenario (details discussed later). The operating model process consisted of three general steps. 1. Define an equilibrium virgin spawning stock biomass (SSB_0), population life history parameters, and fishery characteristics (including reported and unreported catch and tag-releases). 2. Simulate 1000 sets of tag-recapture and catch-at-age observations that reflect the information from step 1 (i.e., if parameter estimation using the observations was performed, the expected values of the observations as estimated by the model would equal the specified values in step 1). The observations are randomly distributed according to the error assumptions defined for the observations and recruitment variability. 3. Calculate a "true" time series of SSB including simulation variability using the information from steps 1 and 2.

The estimation model used the same sets of observations and information used by the operating model with the exception that unreported catch was removed according to scenario and observational data was reweighted. The SSB was estimated under these biased conditions to

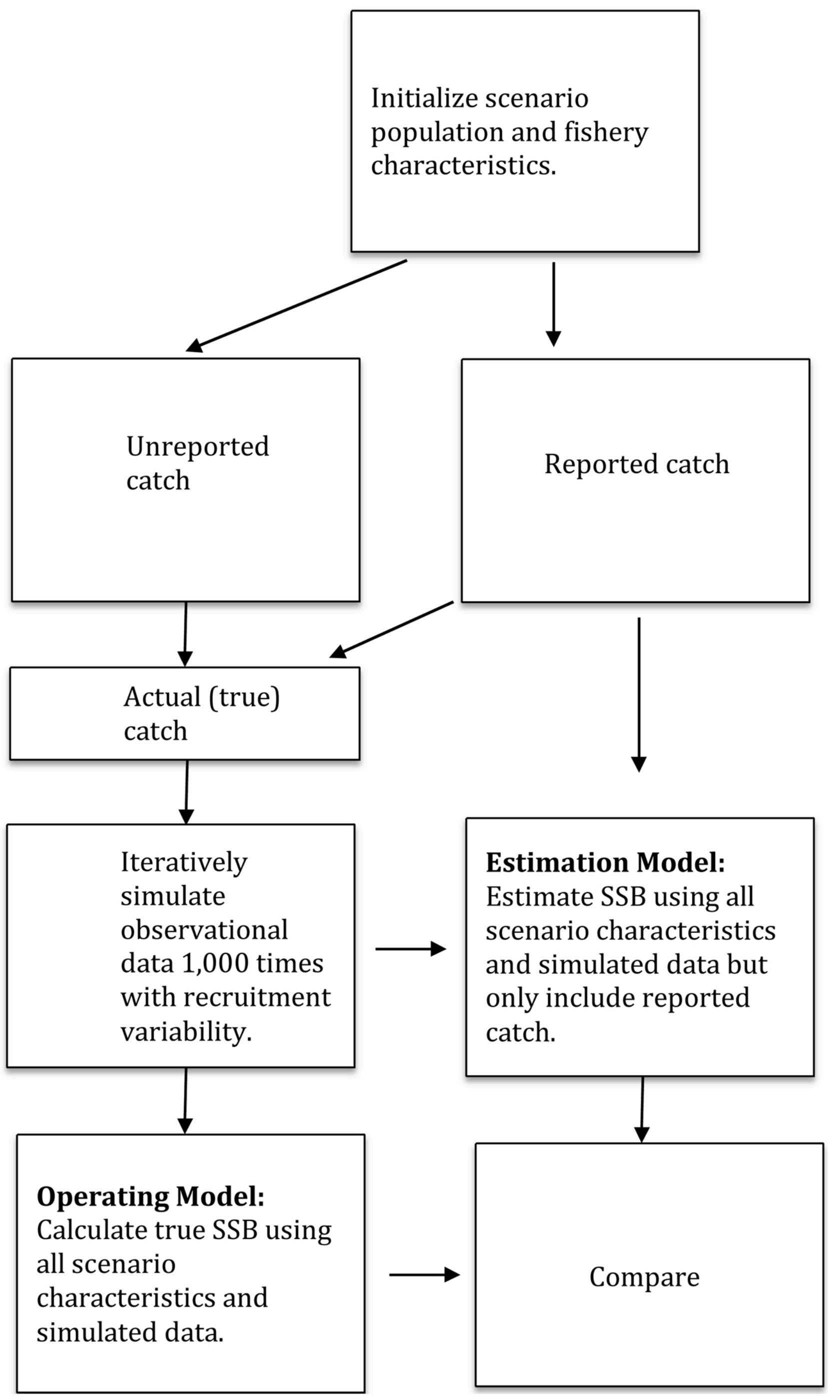

Fig. 2. Model framework used to generate both "true" and "estimated" time series of SSB using an operating model and assessment model respectively; run for different under-reported catch scenarios. CASAL was used for both models and all scenarios converged with a gradient of 0.001. Relevant Patagonian toothfish fishery parameters were estimated from a recent fishery report (CCAMLR, 2018a).

provide an "estimated" time series of SSB including estimation uncertainty to compare against the "true" time series (see supplementary material A for more detail on the CASAL-specific methods).

2.2. population and simulation characteristics

The simulation period for this study was 21 years to reflect a substantial period of under-reported catch and provide adequate time for the stock to respond to fishing. Over this time frame, population characteristics and prescribed fishing fleet dynamics (including catch) were used to simulate observational data necessary for the operating and estimation models. The population was considered to be at unfished equilibrium in the first year, with SSB_0 starting at 100,000 tonnes for all scenarios. Catch was then applied annually over the subsequent 20 years (see catch scenario section below). Observational data was collected in the later period of fishing. In the estimation model, a stock assessment was carried out in final year to estimate the SSB generated by the operating model for the entire 21 period.

The CASAL model used here is similar to the Heard Island and McDonald Islands (HIMI) Patagonian toothfish assessment model initially described by Candy and Constable (2008). The model integrates fishery catch and tag-recapture data (among other information) to estimate SSB. Catch data act as a minimum limit of population numbers, an indicator of productivity, and contributes to estimated fishing mortality. Tag-recapture data is primarily used as an indicator of absolute abundance, although it can inform movement rates. Patagonian toothfish biological parameters were taken from the latest fishery report (CCAMLR, 2018a) for the stock in the Australian Exclusive Economic Zone (EEZ) around HIMI (CCAMLR Division 58.5.2) and are provided in Table 1. Fishing selectivity and data collection details were chosen based on a typical, hypothetical Patagonian toothfish stock. The only free parameters included in the estimation process were SSB_0 (initial recruitment, R_0 was derived internally from SSB_0), and annual recruitment deviations. These population parameters are often the minimal parameters that are estimated in integrated assessments and including selectivity as an additional free parameter did not change the results (see section 3.3).

Two types of observations were simulated, tag-release and recapture-at-length and catch-at-age data. These data types are routinely collected by most established Patagonian toothfish fisheries (CCAMLR, 2018a, 2018b, 2018c, 2018d) and are the dominant types of data (index of abundance and catch composition data) used in integrated assessment models in general (Francis, 2017). The number of fish tagged and released were set annually at 2000 for five years from year 15–19, and roughly corresponds to the number released each year by the HIMI fishery. Tagged fish were available for recapture by the fishery between years 16 and 20. Tag detection and reporting probability were set at one (consistent with toothfish fisheries which often have 100 % observer coverage), tag-release mortality at zero, and tag shedding rate at 0.0084. A single area and stock were defined and tagged fish were simulated to homogeneously mix with the untagged population to ensure mixing assumptions were valid.

The size distribution of fish selected to be tagged and released followed the same distribution as fishing selectivity to ensure unbiased size selection between fish caught and fish tagged (Ziegler, 2013). In addition, the size distribution of recaptured tagged fish in the scanned catch were representative of the fish selected during recapture years, similar to the length-dependent selection model in Tuck et al. (2003).

Between years 11 and 21 (inclusive), 1000 catch-at-age samples were simulated from the catch each year in the operating model by way of proportions in age classes. The expected values of the proportions were passed to the estimation model, but the effective sample size was reduced from 1,000–10 to reduce the weighting on catch-at-age data. Composition data such as catch-at-age is often in conflict with the index of abundance and best practice often results in downweighting the composition information (Francis, 2017). Effectively, this reduces the confidence in the age composition data for the estimation model compared to the operating model to ensure tag-recapture data drives SSB estimation.

Table 1
Population, fishery, and data collection specifics for Patagonian toothfish used in the simulation scenarios.

Parameter	Value	Source
Study Period		
Simulation and assessment period (years)	21 years	
Year of the stock assessment	Year 21	
Population Parameters		
SSB_0 for initializing simulations	100,000 Tonnes	
Age range	1–35 years with plus group	CCAMLR (2018a)
Stock Recruitment Relationship	Beverton-Holt (steepness = 0.75)	CCAMLR (2018a)
Recruitment standard deviation (σ_R)	0.6 (lognormal)	
Natural mortality rate (instantaneous)	0.155	CCAMLR (2018a)
Growth (von-Bertalanffy)	$L_\infty = 1605$mm $k = 0.049$ $t_0 = -3.64$ $CV = 0.131$	CCAMLR (2018a)
Mass at length (mm to tonnes)	$b = 2.59 \times 10^{-12}$ $c = 3.2064$	CCAMLR (2018a)
Maturity (logistic)	$a_{50} = 11$ $a_{95} = 17$	CCAMLR (2018a)
Fishery Information		
Fishing period (years)	Year 2–21 (20)	
Number of fleets	2	
Fleet names	"Legal/reported" and "Illegal/unreported"	
Selectivity (double normal ogive)	$a_1 = 10$ $\sigma_L = 2$ $\sigma_R = 10$	
Catch levels: reported	3,000t	CCAMLR (2018a)
Unreported	10 %, 50 %, 100 %, 150 % of reported catch	
Observations		
Tags		
Length classes for tag observations	300–3050 mm (by 50 mm bins)	
Tag release period (years)	Year 15–20 (5)	
Tag recapture period (years)	Year 16–21 (5)	
Tag number (annual)	2000	
Tag release mortality	0	
Tag shedding rate	0.0084	CCAMLR (2018a)
Tag detection probability	1	
Tag reporting probability	1 (legal fleet), 0 (illegal fleet)	
Proportion of catch scanned for tag	1	
Catch-at-age		
Sampling period (years) Sampling fleet	Year 11–21 (10) Legal	
Sample size for annual catch-at-age	1000 in operating model, 10 in estimation model	
Assumed distribution	Multinomial	
Ageing error	0	
Estimated Parameters		
Recruitment YCS	Lognormal prior: $\mu = 1$, cv = 1 (L: 0.001, U: 100)	
SSB_0	Uniform prior: (L: 5000; U: 500,000)	

2.3. catch scenarios

The fishery in the operating model was split into two fleets to simulate removals of fish from the population throughout the fishing period. These two fleets are nominally distinguished here by a fleet which reports its catch and recaptured tagged fish exactly (legal fleet) and a fleet that does not report any catch or recaptured tagged fish (illegal fleet). The legal fleet released tagged fish and collected catch-at-age data through its catch. The illegal fleet did not tag or age fish, but it did remove tagged fish from the population in proportion to its catch. Observation data were therefore only provided to the estimation model through the legal fleet but reflected the state of the population under fishing pressure from both fleets. There was therefore a potential conflict in the estimation model between the information about population changes provided by the tag-recapture and catch-at-age composition data, and the catch that has led to these changes.

Employing simulation analyses allowed us to ensure that all the assumptions for a tag-recapture study were met except the one we wanted to examine. Administering the tag-releases and recaptures through the legal fleet, combined with the fixed tag-related parameters (tag-induced mortality, tag-reporting, tag-detection, tag-shedding), and tagged fish being representative of the population, lead to all assumptions required for open population tag-recapture studies (see Pine et al. (2003) for a review of required assumptions) to be met except that tagged fish caught illegally were not reported. The combined impact of unreported tagged and untagged fish caught illegally on an integrated assessment outcome could therefore be isolated and examined here.

Illegal catch scenarios were chosen based on their realistic probability for toothfish specifically and other fisheries vulnerable to IUU fishing in general. IUU catch reconstructions for Patagonian toothfish estimate that IUU fishing peaked in the initial phase of four major fisheries in the sub-Antarctic (Fig. 1). Estimated IUU catch was also highly variable during this initial phase. In the later phase of these fisheries, IUU catch was either estimated to reduce to zero or was no longer estimated. Linearly trending IUU catch scenarios were applied here in order to generalize scenarios, to simplify analysis, and to be consistent with scenarios investigated in previous studies (Omori et al., 2016; Rudd and Branch, 2017). Three types of linear trends in IUU catch were applied, namely constant, decreasing to zero, and increasing from zero.

The magnitudes of illegal catch were defined as a proportion of legal catch. The current total allowable catch for Patagonian toothfish at HIMI is approximately 3000 tonnes based on a maximum sustainable yield estimate with a target SSB of 50 % SSB_0 (CCAMLR, 2018a). In all scenarios examined here, a constant reported catch of 3000 tonnes for each year of the fishery period was applied. A combination of the three types of IUU catch trends and four cumulative levels of 10 %, 50 %, 100 %, and 150 % of the reported catch were considered in a total of 12 scenarios

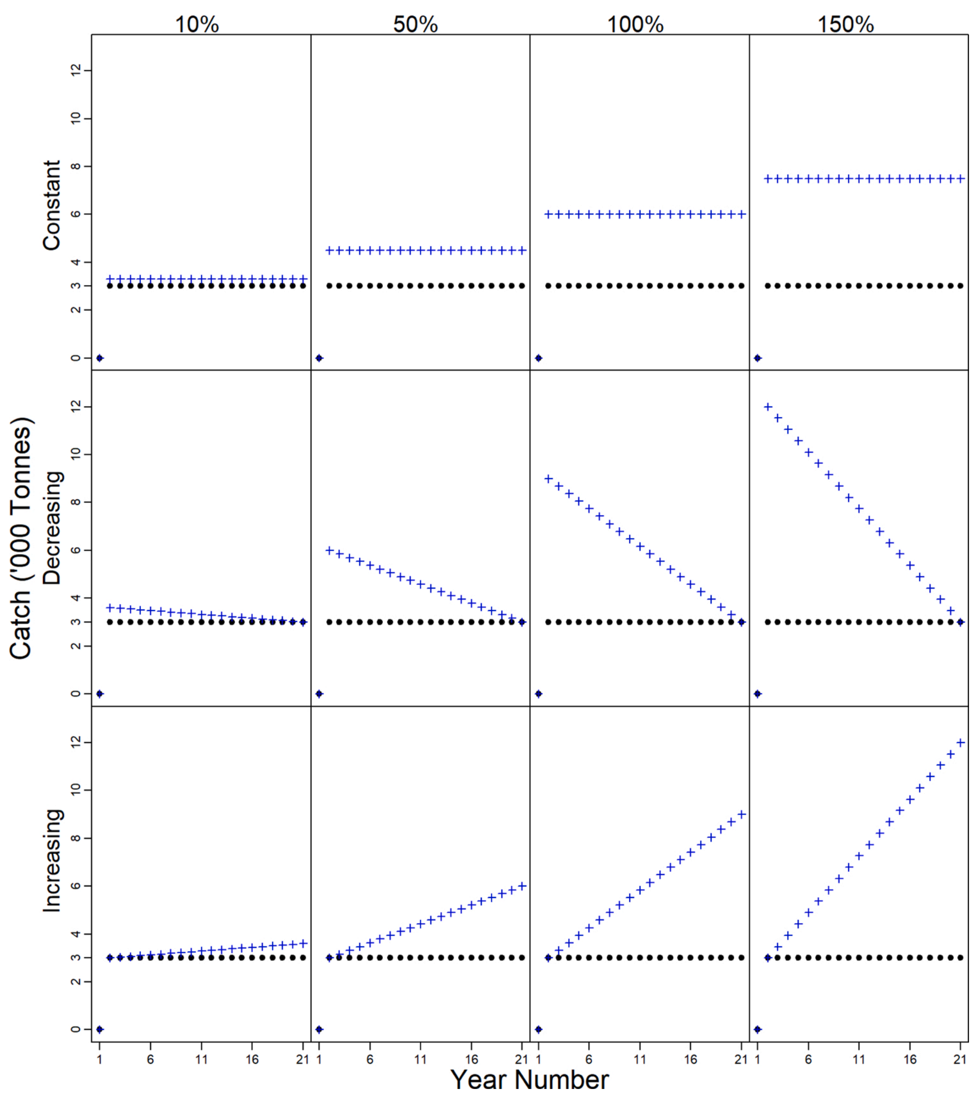

Fig. 3. Catch scenarios used in the simulations with reported catch (black dots) and actual catch (reported and unreported catch, blue crosses). The rows represent three types of unreported catch (constant, decreasing, and increasing), while the columns represent four magnitudes (10 %, 50 %, 100 %, 150 %) of unreported catch for a total of twelve scenarios. No fishing occurred in the initial year to initiate the population at equilibrium.

(Fig. 3).

Given the chosen catch scenarios, along with starting each scenario at the same unfished SSB_0, we would expect scenarios with higher catches in the operating model to experience more depletion than those with lower catches. The decision to start each scenario at the same SSB_0 was intended to allow for the most like-with-like comparison across trend types. If scenarios started at different SSB_0 to, for example, ensure constant depletion by the assessment year in the operating model, the impact of trend type on depletion would be less clear. However, by not ensuring constant depletion, there is a potential for the results to be confounded by the combined impact of catch magnitude and depletion, particularly considering the stock-recruitment relationship in this example. To address this, we ran additional scenarios as sensitivities where different values of SSB_0 were chosen to ensure constant depletion across all scenarios (see section 3.3).

2.4. scenario evaluation

Under each scenario, the raw time series of SSB from the operating and estimation models were compared as well as individual characteristics (virgin, current, and status) of the SSB time series important for management decisions, where

SSB_0 = virgin SSB at equilibrium in year 1; before fishing commenced

$SSB_{current}$ = current SSB during the assessment in year 21

$SSB_{status} = (SSB_{current}/SSB_0)$

These values were compared against management target ($SSB_{status} = 0.5$) and limit ($SSB_{status} = 0.2$) reference points used for CCAMLR toothfish fisheries (Constable et al., 2000).

The difference between the true SSB characteristics simulated by the operating model and those estimated by the estimation model were also compared using a relative error statistic:

$$E_{SSB} = \frac{SSB_{est} - SSB_{true}}{SSB_{true}}$$

where a relative error (E_{SSB}) of 0.5 indicates the assessment overestimated SSB by 50 % while a value of -0.5 indicates it was underestimated by 50 %. The relative error of SSB_0, $SSB_{current}$, and SSB_{status} were all examined as they are each useful in determining the assessment bias and are important for management concerns.

2.5. Lincoln-Petersen comparison

Estimates from a Lincoln-Petersen (LP) estimator were compared to the model output from CASAL to act as a check on the influence of tagging data on the results. The LP estimator is a simple model that uses only tag-recapture data and catch to estimate population abundance either in terms of numbers or biomass (Seber, 1982). The tag-recapture data was extracted from the simulations and used in a LP estimator to calculate vulnerable population numbers and compare them to vulnerable population numbers output from CASAL. Vulnerable numbers represent the number of fish in the population that are vulnerable to being selected by the fishery, as defined by the selectivity ogive. The comparison was limited to vulnerable numbers instead of SSB for two reasons. 1. The LP estimator cannot be used to calculate SSB since it cannot incorporate age structure and therefore has no information about mature biomass. 2. The LP estimator is limited by how representative the simulated tag-recapture data is. On its own, this data only represents vulnerable population numbers (number of fish available for selection by the fishery), not SSB since only fish vulnerable to selection by the fishery can be recaptured. The LP results were therefore compared to vulnerable population numbers calculated by the operating model and those estimated by the estimation model.

The comparison allowed us to investigate the cause of biases between the operating and estimation model results. It is often difficult to identify what source or sources of information are driving potential biases in integrated assessments and the LP estimator was useful for determining the potential influence of the tag-recapture data specifically. The LP estimator was chosen over more complicated tag-recapture models to check if the simplest model possible could reproduce the characteristics of the CASAL output. If a simple model such the LP estimator detected biases in the tag-recapture data that correspond to biases from the CASAL output, then the tag-recapture data were likely the cause for biases in abundance estimates like SSB. A more complicated tag-recapture model may provide similar results but identifying the root cause of bias due to the tag-recapture data instead of another model attribute would be difficult. Note that the LP estimator as it is used here does not act as an alternative method of estimation. It only acts as a way to check for the cause of bias. The tag-recapture data used to calculate the LP estimator originated from the CASAL model, which has known biases (catch). Thus, the LP cannot be an independent alternative to estimate abundance.

The LP estimate incorporated the total annual number of tag releases and subsequent recaptures, natural mortality, and tag-shedding rate (Table 1) to estimate vulnerable population numbers during the recapture years 16–21 (see supplementary material B for specific LP formulation). The numbers-by-length of recaptured tagged fish were aggregated over length classes to achieve a total number of fish recaptured in each recapture year as the LP estimate is not length-structured while the simulated tag-recapture data are. We compared the relative error in vulnerable numbers from CASAL ($E_{N\text{-}CASAL}$) between the operating and estimation model to the relative error in vulnerable numbers from the LP estimator ($E_{N\text{-}LP}$) using the tag-recapture data from each model.

Since the tag-recapture data was not changed when passed to the estimation model in the main results (Fig. 2), directly using that data to calculate LP estimates to represent the estimation model would yield no differences to those from the operating model. To address this, different sets of tag-recapture data were used to represent the operating and estimation models in the LP formulation. A set of data was generated to reflect expected recaptures from the actual catch to represent the operating model and another set was generated using only the reported catch to represent the estimation model. The LP estimator was then applied to the two sets of tag-recapture data to achieve “true” and “estimated” vulnerable population numbers for each scenario. E_N was then calculated based on these estimates. Notably, the model framework was repeated using the different sets of tag-recapture data and there were minimal changes (results not shown) to the SSB time series presented as the main results.

3. Results

3.1. integrated assessment

The general trajectories of the true and estimated SSB over the study period illustrate the differences between the true and estimated states of the population under the 12 scenarios of under-reported catch (Fig. 4). From an SSB_0 of around 100,000 tonnes in year 1, the true SSB declined gradually after fishing commenced in all scenarios. As expected, the decrease was faster where unreported catch and hence the actual catch was higher. When actual catch was 100 % or 150 % greater than reported catch, the true SSB dropped below 50 % SSB_0 (the target reference point for this stock) by the assessment year. In none of the scenarios did the true or estimated SSB drop below the limit reference point of 20 % SSB_0 within the simulation period.

In all scenario trend types (rows in Fig. 4), the variability in both the true and estimated SSB trajectories decreased as severity of under-reported catch increased (columns in Fig. 4). This is likely because the amount of observational data collected across scenarios was the same regardless of catch, resulting in more data collected relative to

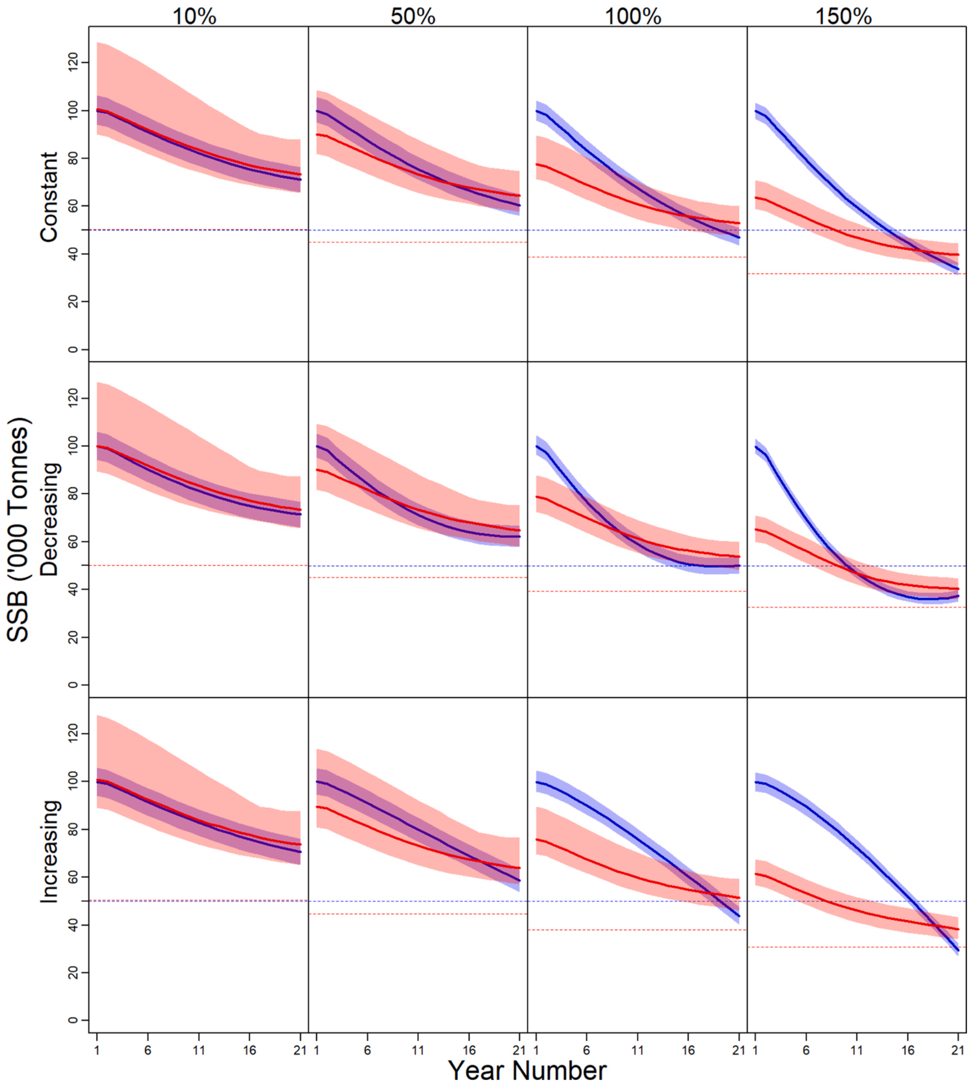

Fig. 4. Spawning stock biomass trajectories corresponding to the 12 catch mis-specification scenarios (Fig. 3) over the study period. Blue lines and surrounding shading are the median and 95 % simulation intervals of the true SSB from the operating model. Red lines and surrounding shading are the median and 95 % confidence intervals of the estimated SSB from the estimation model. Dashed lines represent the 50 % SSB_0 level for their respective colors/models.

population size for scenarios that had a larger impact on biomass (i.e., larger catch). There was also greater variability in the estimated SSB than the true SSB for all scenarios. This was due to the reduced confidence assigned to the catch-at-age data in the estimation model. When there is no sample size reduction for catch-at-age data, the variability between models is comparable (see supplementary material C.4; Francis, 2017).

For the lowest magnitude of unreported catch (10 %), the median SSB was well estimated by the estimation model. This implies that investigating magnitudes lower than 10 % would be trivial. As the magnitude of under-reporting increased, the estimation model estimates became progressively worse and failed to capture key features of the SSB trajectory. In particular, estimates of SSB_0 worsened as magnitude of under-reported catch increased. For example, SSB_0 was estimated at roughly 60,000 tonnes while in reality it was 100,000 tonnes when under-reported catch was at 150 %. This was the case across all three trend types of unreported catch. While the true SSB fell below the true target reference point at some stage during the simulation in some scenarios, the estimated SSB did not fall below the estimated target reference point. The slope of SSB was consistently underestimated and consequently the level of depletion (SSB_{status}) was also underestimated (Fig. 5).

When the magnitude of under-reported catch was at least 100 % greater than reported catch, SSB fell below the target reference point of 50 % SSB_0. However, the estimation models failed to identify this. For example, when under-reported catch was 150 % greater than reported catch for all catch trend types, the true SSB_{status} dropped to 35 % by year 21, while it was estimated at 60 % by the estimation model. This result was consistent regardless of the type of trend in under-reported catch. In the first decade of fishing, however, estimates of SSB_{status} in all scenarios of increasing under-reported catch matched the true values well. This was not the case for the other trend types of under-reported catch. For decreasing and constant types, the depletion level was underestimated over the entirety of the study period.

Trends in true SSB_{status} varied markedly across scenarios while the trends in the estimated SSB_{status} did not. True SSB_{status} decreased to a lower point over time as the magnitude of under-reported catch increased in all scenarios. However, the rate of decrease was affected by the trend in under-reporting. For example, SSB_{status} trajectories for constant and decreasing trends were concave up, meaning the rate of depletion was decelerating. This feature was more pronounced for the decreasing trend type than for the constant type. Conversely, the rate of depletion was accelerating when unreported catch trended upwards. These characteristics of the true SSB are not unexpected, however they

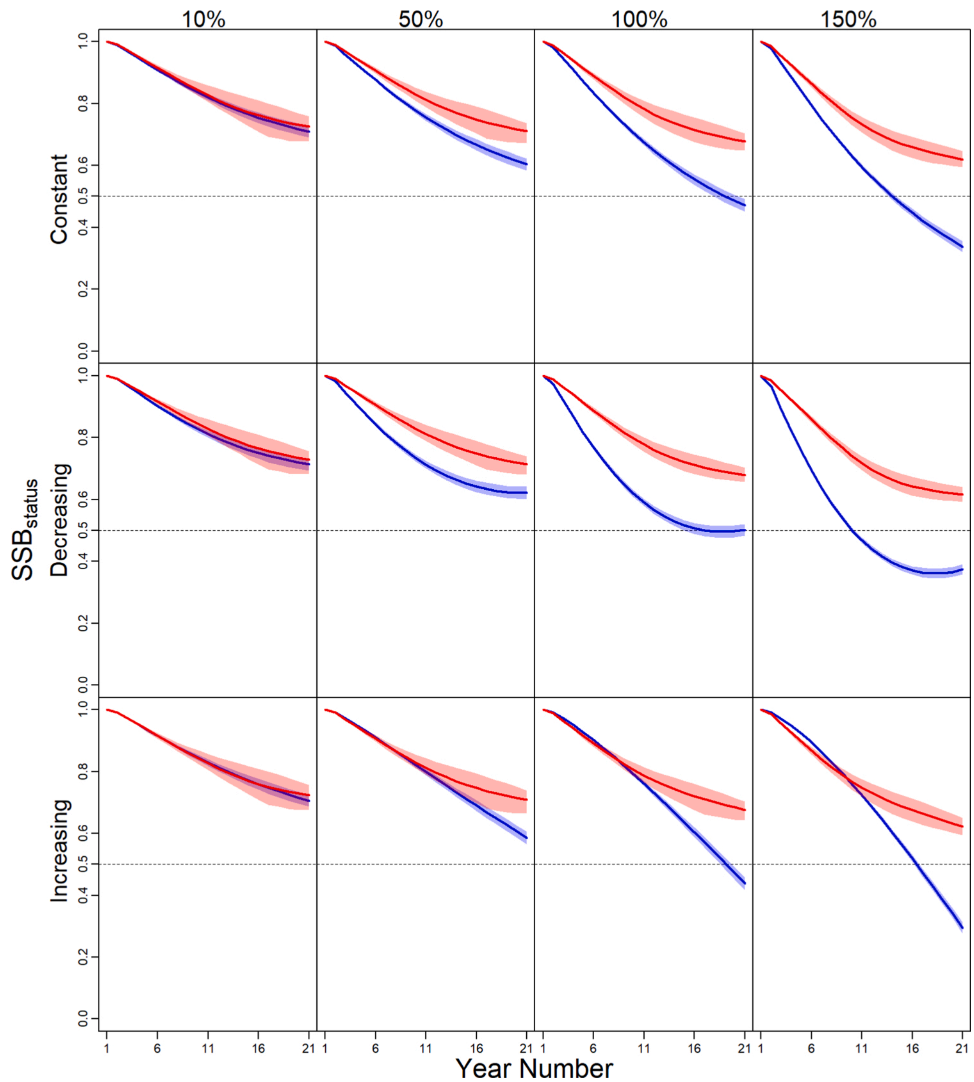

Fig. 5. Spawning stock biomass status (current/virgin) evolution throughout the study period for the operating model (true; median blue line and 95 % simulation interval shading) and the estimation model (estimated; median red line and 95 % confidence interval shading) under the different catch scenarios in Fig. 3. Horizontal dashed black line indicates 50 % management reference point.

were not captured by the estimation model. In fact, all estimation model estimates of SSB_{status} appeared similar regardless of scenario. The only exception was for 150 % under-reporting, where the estimated SSB_{status} was slightly lower for all three trend types by the time of the assessment year compared to 10 %, 50 %, or 100 % under-reporting.

These biases can be more formally quantified using the relative error statistic. The relative error between the models for the current, virgin, and status features of the SSB is provided in Fig. 6. Notably, the estimation model estimated SSB most accurately for the current period of the study as opposed to the initial period for all scenarios. This is particularly evident for the decreasing under-report of 150 % scenario (Fig. 6). Presumably, this reflects the ability of tag-recapture data to perform better as an index of abundance during the period in which tag-recapture data is being collected. Shifting the tag-recapture effort to the start of the study reveals more accurate initial biomass estimates and less accurate current estimates (results not shown). However, collecting tag-recapture data during the initial phase of a fishery and not the current phase is unrealistic in practice.

As the percent under-reported catch increased, the absolute value of relative error in all three SSB metrics increased across all scenarios (Fig. 6). The estimation model increasingly overestimated $SSB_{current}$ and underestimated SSB_0 as the percent under-reported catch increased from 10 % to 150 %. SSB_{status} was overestimated by a larger degree than either $SSB_{current}$ or SSB_{status} as it is a ratio between the two. The most pronounced bias occurred for SSB_{status} at 150 % increasing under-reporting where the estimate was more than double the true value. This indicates that the estimation model captured less than half of the true SSB depletion. There was a similar, but less severe, bias for the other trend types of under-reported catch occurring at the same magnitude. The only exception was that $SSB_{current}$ was consistently better estimated for the decreasing trend type compared to the others.

3.2. Lincoln-Petersen comparison

The relative error between the operating and estimation model vulnerable population numbers from CASAL ($E_{N\text{-}CASAL}$) was compared to the relative error in vulnerable numbers using a Lincoln-Petersen (LP) estimate ($E_{N\text{-}LP}$) (Seber, 1982) during the tag-recapture phase (years 16–21) (Fig. 7). The trends in relative error estimates generated by the LP estimate were overall similar to those from CASAL for each scenario. At 10 % under-reported catch, both estimates reported zero error. As under-reported catch magnitude increased to 150 % for all scenarios, the absolute values of error matched less, however the trends were still comparable. In year 17, $E_{N\text{-}LP}$ reported no bias for all scenarios, but as

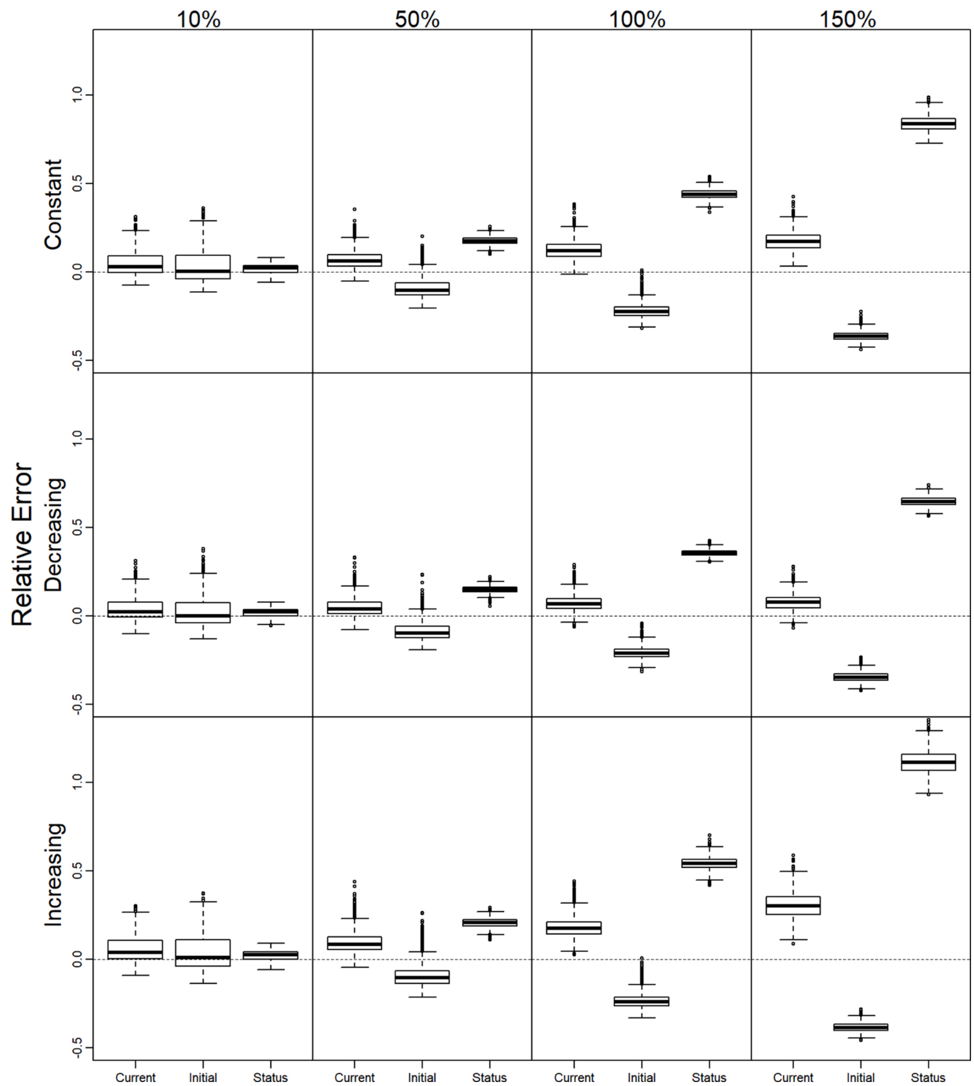

Fig. 6. Relative error between the operating model and estimation model for virgin (pre-fishing, year 1), current (assessment year 21), and status (current/virgin) of SSB across catch scenarios represented as boxplots.

more cohorts of tagged fish were recaptured in later years, the bias increased. This was not unexpected as with more recapture data, the LP estimator should be able to detect greater discrepancies between data sets.

3.3. sensitivity analyses

Several sensitivity analyses were undertaken to identify the impact of data collection assumptions on the results (supplementary material C). Constant annual tag-releases applied through different levels of catch (as was performed for the main results) implies that the quantity of fish tagged and released is not scaled with catch. This is not true in all real-world cases where sometimes the number of fish required to be tagged by the fishery scales with their catch. Therefore, sensitivities to test the impact of tagging effort were performed. Increasing or decreasing the number of years tag-recapture data was collected as well as increasing the number of fish that were tagged each year had little influence on the results (supplementary material C.1–C.3). This suggests that some minimum absolute number of tag-releases (instead of a tag-release rate proportional to catch) may be suitable to produce biomass estimates, but properly investigating this concept is outside the scope of this study. As another sensitivity, consistent age composition sample sizes were maintained between the operating and estimation model at 1,000. Unlike changing tagging effort, this sensitivity did change the results (supplementary material C.4). The uncertainty surrounding all biomass estimates decreased and the accuracy of SSB_{status} for 100 % and 150 % under-reported catch magnitudes increased. The structure of the under-reported catch scenarios was also changed from linear to a step change (supplementary material C.5). Even under these dramatically different catch conditions, the results changed very little. Finally, SSB_0 was changed in each scenario in the operating model to ensure constant depletion (SSB_{status}) was reached by the assessment year. Depletion was targeted at 50 %, 70 %, and 30 % (supplementary material C.6–C.8) to determine if the results were impacted by confounding depletion with catch. Although the discrepancies between true and estimated SSB were scaled according to the target depletion level, constant depletion had no qualitative impact on the results. SSB_{status} was increasingly under-estimated with increasing illegal catch over all trend types, similar to the main results using non-constant depletion.

In addition, a sensitivity analysis was performed in which fishing selectivity was estimated as an additional free parameter within the estimation model (supplementary material C.9). This sensitivity therefore had three free parameters in total, including the two standard free parameters, SSB_0 and annual recruitment deviations. The uncertainty in estimated SSB_{status} increased slightly in all scenarios, but the results remained largely unchanged.

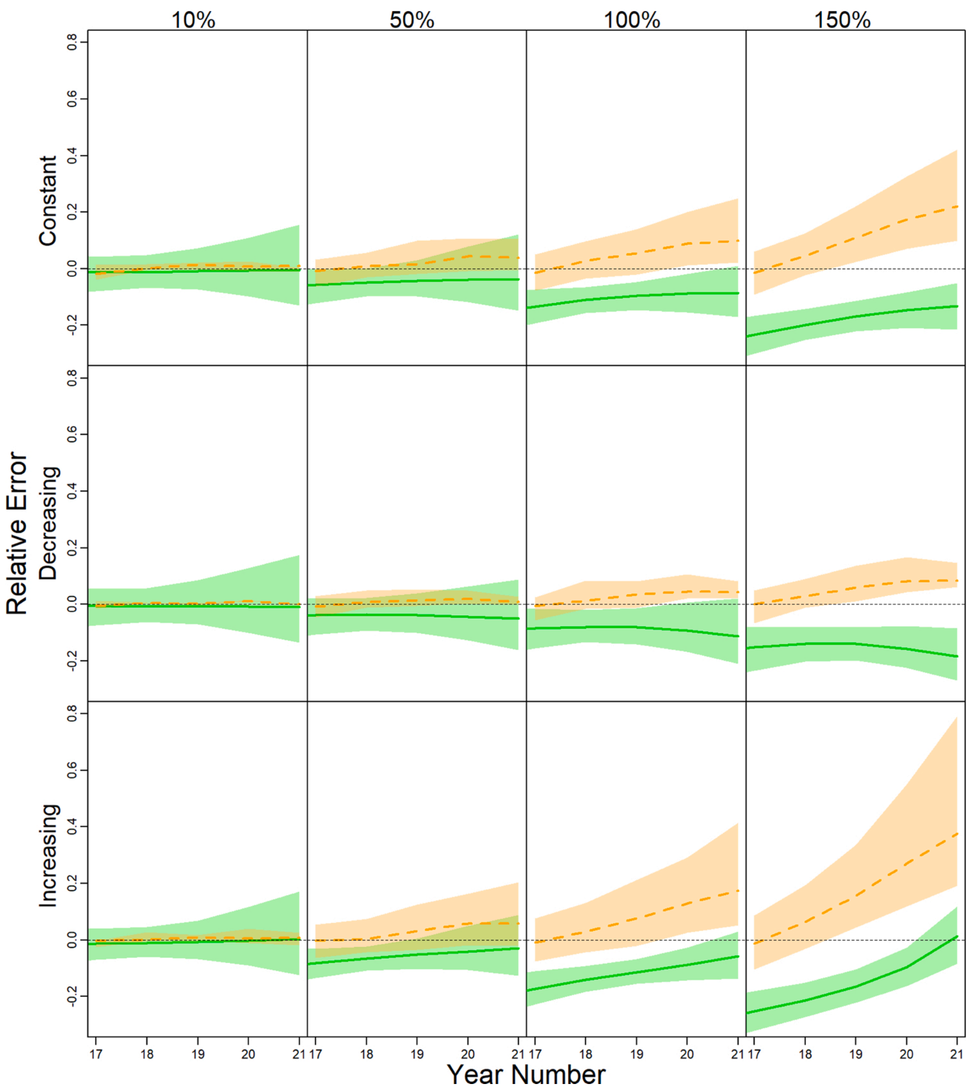

Fig. 7. Relative error between the operating and estimation models from CASAL estimated vulnerable population numbers ($E_{N\text{-}CASAL}$, median solid green lines and 95 % confidence interval shading) and mean (averaged over tag-recapture cohorts) Lincoln-Petersen (LP) estimated vulnerable population numbers ($E_{N\text{-}LP}$, median dashed orange lines and 95 % confidence interval shading). Comparison is limited to years 16–21 as these were the only years tag-recapture data was available and for which the LP estimate could be calculated.

4. Discussion

The results demonstrate that spawning stock biomass estimates (SSB_0 through to $SSB_{current}$) from a tag-based integrated assessment model are likely to be biased if there is substantial under-reported catch and removal of tagged fish. Furthermore, estimates of relative biomass (SSB_{status}) are also likely to be biased even if there is moderate under-reported catch. These biases occur regardless of the temporal trend in reporting rate and would result in an overall unconservative estimate of depletion if not accounted for. In contrast, Rudd and Branch (2017) and Omori et al. (2016) found biomass depletion estimates were only biased by increasing or decreasing trends in misreported catch and were unbiased by constant levels of under-reporting. This contrast is important considering constant trends in under-reporting are the most common on average among the world fisheries (Pauly and Zeller, 2016) but not all those fisheries are assessed with surplus production models. A key difference between the methods in Van Beveren et al. (2020) and those employed here was that the tag-recapture index of abundance used in this study was biased in the estimation model due to unreported catch whereas the index of abundance (total egg production) remained unbiased when implementing harvest control rules with unreported catch in Van Beveren et al. (2020). The difference allowed for the tag-recapture data to be specifically examined as a potential source of information leading to the misestimation of SSB, whereas determining the impact of the index of abundance on the estimation performance of SSB was not a focus of Van Beveren et al. (2020). This study is more likely to be applicable to fisheries that rely on integrated (as opposed to surplus production) models that incorporate tag-recapture data (as opposed to other indices of abundance, e.g., catch rates or egg production) as an indicator of abundance for assessment.

The magnitudes and trend types of actual catch compared to reported catch proposed here are similar to those estimated for worldwide catch reconstructions within the 19 maritime statistical areas defined by the Food and Agriculture Organization (Pauly and Zeller, 2016). The catch scenarios also represent the true history of most Patagonian toothfish fisheries where both legal and illegal fishing have occurred simultaneously throughout the sub-Antarctic (Agnew, 2000; Agnew and Kirkwood, 2005). The fundamental characteristics of these catch reconstructions were considered to develop idealized catch scenarios for this study. We found that once under-reported catch exceeded 100 % of reported catch (IUU catch exceeds legal catch) the true SSB fell below the 50 % target reference point but this was undetected by the estimation model (Fig. 5). These results remained largely unchanged under the sensitivities we considered that varied the tag-release quantity, number of release/recapture years, true depletion level at the time of assessment, the structure of under-reported catch history, and estimating fishing selectivity in addition to SSB_0 and recruitment deviations (see supplementary material C). The consequence of this in a real assessment would be an underestimate of fishing impact on the population and subsequently a failure to take management action to maintain the population at the chosen target level. This indicates that tag-based integrated assessments are not robust to IUU fishing.

Although the trends in true SSB simulated by the operating model varied markedly among scenarios, the trends in estimated SSB remained largely unchanged. The resulting bias between true and estimated SSB status was therefore intensified in proportion to the magnitude of under-reported catch (Fig. 6). Only at 150 % under-reported catch did the estimation model start to detect a greater decrease in SSB_{status} (Fig. 5). This result was exaggerated when age composition weighting was higher (see Supplementary Material C.4). This marginal improvement was still far from the true value but may mark the point at which bias in tag-recapture data begins to take effect in an estimation model conditional on a small sample size of age composition data. If this is the case, a substantial discrepancy between catch and tag-recapture data may be required for a tag-based assessment model to detect changes in biomass depletion rate. The precise nature of this discrepancy warrants further investigation, especially considering the potential impact of conflicting data sources on integrated assessments (Francis, 2017, 2011).

The trends in misestimation of vulnerable population numbers during the last five years of the study were detected through vulnerable population numbers estimated by the LP estimator that relied solely on the simulated tag-recapture data and catch (Fig. 7). Similar to the CASAL-estimated vulnerable numbers, the LP estimator reported greater discrepancies between the operating and estimation model with greater under-reported catch magnitude. It was unsurprising that the relative error generated by the LP estimator did not exactly match the relative error generated by CASAL as the LP estimates were not moderated by the population structure, recruitments, deaths, or catch-at-age data before or during the tag-recapture period.

Tag-recapture data was the primary source of information used by the estimation model to inform SSB_0, a key component for calculating SSB_{status}. Estimates of SSB were more erroneous the further away from the tag-recapture period and the closer to the pre-fishing year (year 1) they were, leading to large underestimates of SSB_0. The tag-recapture data likely provided less and less reliable information closer to the pre-fishing year, contributing to increasingly biased estimates of SSB toward that year. The fact that the trend in LP estimates of relative error generally concurred with those from CASAL supports the notion that the tag-recapture data was responsible for mis-leading estimates of vulnerable numbers and subsequently SSB. Tag-recapture methods are generally considered superior to using catch rates to inform abundance (Walters and Martell, 2002) but here they have failed to accurately represent it in the integrated approach. This is contrary to the results expected by Welsford and Ziegler (2013) for *Dissostichus* spp. that suggest estimates would be unbiased even if IUU fishing has occurred, as long as IUU fishing removed tagged and untagged fish in the same proportion as legal fishing. Although this condition was met in this study, the combination of using under-reported catch and failing to incorporate all recaptured tagged fish misled the estimation model to substantially underestimate depletion.

The impact of under-reported catch on an integrated assessment model is difficult to predict given their dependence on many sources of information. Many fisheries throughout the world are assessed using integrated stock assessment models like the one used in this study (Maunder and Punt, 2013). Tagging studies have been incorporated into such models to provide direct information about abundance among other important quantities such as fishing and natural mortality rates (Ainsworth and Pitcher, 2005; Hoenig et al., 1998; Polacheck et al., 2006; Pollock et al., 2002a; Ricker, 1975). Tag-recapture data continue to be more and more common in modern stock assessments due in part to improved computational capability (Hoyle et al., 2020) and the capacity to incorporate tagging information is listed as an essential component in the next generation stock assessment model (Punt et al., 2020). For assessments reliant on such data, it is important to understand that the impact of misreported catch may not be intuitive. Refraining from thoroughly investigating this impact using simulation analysis involving the stock assessment model used in practice and relying instead on assumed theoretical outcomes or published results from dissimilar models, could have poor consequences. For fisheries reliant on integrated assessments that use tag-recapture data to estimate abundance, it is important to reconstruct IUU catch as accurately as possible to ensure that reliable estimates of SSB_{status} are available to guide sustainable catch limits for the fishery. Specifically, this study demonstrates that it is important to accurately estimate the absolute quantity of IUU catch if it has the potential to be substantial. Although the different trends in under-reported catch had slightly different impacts on depletion estimates, it was the overall quantity of catch that caused the most bias in this study. This is an important consideration for other high value species traditionally vulnerable to IUU fishing that depend on tag-recapture data such as several species of tuna. Southern bluefin tuna for example experienced substantial levels of IUU fishing for decades which contributed to their severe, undetected, decline in biomass (Polacheck, 2012).

For most fisheries where IUU fishing is taking place, estimating IUU catch accurately is difficult. In these cases, it may be useful to establish a minimum acceptable criterion for the certainty surrounding IUU catch reconstructions. This is similar to the criteria many fisheries employ regarding the probability that biomass ratios exceed management thresholds (Caddy and Mahon, 1995; Gabriel and Mace, 1999). For the case study examined here, a catch reconstruction criterion could be adopted in a similar fashion. For example, we found that when actual catch exceeds reported catch by 50 %, there is a roughly 25 % positive bias in SSB_{status} (Fig. 6). If a positive (unconservative) bias of 25 % is unacceptable for management objectives, then managers would need to be certain that IUU catch is not underestimated by more than 50 %. Pauly and Zeller (2016) describe a practical way of quantifying uncertainty in catch reconstructions using a scoring system based on the quality attributed to catch time series data that would be useful for this procedure. Alternatively, the assessment could be performed with both extremely low and extremely high values of possible IUU catch and both results used to consider management decisions. This technique may be most appropriate for fisheries where estimating IUU catch accurately is impossible. Provided catch is typically considered a non-estimable data source in an assessment model (although there have been modeling techniques developed that treat catch as censored data (Hammond and Trenkel, 2005)), the two approaches suggested above are practical ways IUU catch uncertainty can be addressed in integrated assessments.

Another important consideration is how to treat tag-recapture data collected during periods for which IUU fishing has occurred. This study indicates that tag-recapture data cannot be relied on as an unbiased index of abundance when IUU fishing is removing tagged fish without reporting them. This is the case even if tagged and untagged fish are removed in the same manner from IUU fishing as authorized fishing. Therefore, improving the representation of IUU tag-recaptures alongside IUU catch reconstructions is important. Fisheries for which tag-reporting has been an issue among authorized fishers have adopted high rewards for returning tags (Pollock et al., 2002b, 2001; Sackett et al., 2018) among other strategies to combat this issue. However, this approach would only incentivize the return of tags retrieved through IUU fishing if the reward was worth more than the fish itself and anonymity was guaranteed, an unlikely scenario. If directly improving the representation of IUU tag-recaptures is unlikely, a potential way to account for them is to adjust the tag-detection probability in conjunction with estimates of reconstructed IUU catch within the assessment model. The tag-detection probability specifies the proportion of tagged fish that are detected in the recapture process. This approach has been applied in similar situations where tag-reporting is incomplete and tag-detection rates require adjusting to estimate abundance (Hearn et al., 1999). Similar adjustments have been made to account for inter-vessel and temporal variability in tag-detection probability in the Ross Sea region (Mormede and Dunn, 2013).

CASAL and other contemporary integrated stock assessments that can accept tag-recapture data have the option to adjust the tag-detection probability. For example, if IUU catch was estimated to be equal to reported catch, tag-detection probability was 100 % for reported catch, and the sum of reported and reconstructed IUU catch was entered into the assessment model, it would be reasonable to decrease the tag-detection rate by half to account for the unreported removal of tagged fish. The model would implement this adjustment by doubling the effective number of recaptured tagged fish. Care would be required in estimating IUU catch, however, as the accuracy with which adjusting tag-detection probability depends on it. Furthermore, such an adjustment would likely increase accuracy in estimated abundance during the tag-recapture period but may have unintuitive impacts on abundance estimates outside that period (namely SSB_0). Simulation analysis involving adjusting tag-detection rates would be required to evaluate the effectiveness of compensating for unreported tagged fish removal in this manner.

Overall, we found that tag-recapture data fail to guide an assessment to accurately estimate biomass when catch and tag-recaptures are under-reported in a wide range of catch scenarios. The results are applicable to all fisheries that use integrated assessments that incorporate tag-based indices of abundance where there has been IUU fishing. Furthermore, this study has demonstrated that the impacts of under-reported catch on tag-based integrated assessments are different to that of surplus production models (Omori et al., 2016; Rudd and Branch, 2017). Applying the findings from surplus production models to the integrated context would result in management decisions that are likely not precautionary. Although Van Beveren et al. (2020) employed an integrated approach, the data used to estimate abundance accurately reflected catch uncertainty, something not available to cases such as that investigated here that have an unavoidable conflict between catch and abundance data. If an integrated tag-recapture based model is used to assess a fish stock, simulation analysis similar to that undertaken here should be carried out to allow complete and adequate evaluation of the impact catch and tag-recapture biases have on metrics used to set catch limits. Investigating the impact of IUU fishing using integrated assessments that rely on other types of abundance indicator data (e.g., catch rates) and the role that data plays in potential misestimation would also be useful as tag-recapture data is not available in many fisheries. Future work could investigate incorporating harvest control rules in simulations (management strategy evaluation) as this would be important in determining the robustness of different management strategies to catch and tag-related biases over an extended period.

CRediT authorship contribution statement

Brett Stacy: Conceptualization, Formal analysis, Software, Visualization, Writing - original draft. **Paul Burch:** Supervision, Writing - review & editing, Conceptualization. **Philippe E. Ziegler:** Software, Writing - review & editing. **Katherine A. Cresswell:** Writing - review & editing, Visualization. **Klaas Hartmann:** Supervision, Writing - review & editing. **Richard M. Hillary:** Software, Methodology.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We would like to thank Andre E. Punt for valuable comments to an earlier version of this paper. We are also grateful to the two reviewers of the journal, whose suggestions improved the manuscript. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.fishres.2021.106098.

References

Agnew, D.J., 2000. The illegal and unregulated fishery for toothfish in the Southern Ocean, and the CCAMLR catch documentation scheme. Mar. Policy 24, 361–374.
Agnew, D.J., Kirkwood, G.P., 2005. A statistical method for estimating the level of IUU fishing: application to CCAMLR Subarea 48.3. Ccamlr Sci 12, 119–141.
Agnew, D.J., Pearce, J., Pramod, G., Peatman, T., Watson, R., Beddington, J.R., Pitcher, T.J., 2009. Estimating the worldwide extent of illegal fishing. PLoS One 4, e4570.
Ainsworth, C.H., Pitcher, T.J., 2005. Estimating illegal, unreported and unregulated catch in British Columbia's marine fisheries. Fish. Res. 75, 40–55. https://doi.org/10.1016/j.fishres.2005.05.003.
Beare, D.J., Needle, C.L., Burns, F., Reid, D.G., 2005. Using survey data independently from commercial data in stock assessment: an example using haddock in ICES Division VIa. ICES J. Mar. Sci. 62, 996–1005.
Borges, L., Rogan, E., Officer, R., 2005. Discarding by the demersal fishery in the waters around Ireland. Fish. Res. 76, 1–13.
Bull, B., Francis, R.I.C.C., Dunn, A., McKenzie, A., Gilbert, D.J., Smith, M.H., Bian, R., Fu, D., 2012. CASAL (C++ algorithmic stock assessment laboratory): CASAL User Manual v2.30-2012/03/21. NIWA Tech. Rep. 135, 280 pp.
Caddy, J.F., Mahon, R., 1995. Reference Points for Fisheries Management. Food and Agriculture Organization of the United Nations Rome.
Candy, S.G., Constable, A.J., 2008. An integrated stock assessment for the Patagonian toothfish (Dissostichus eleginoides) for the Heard and McDonald Islands using CASAL. CCAMLR Sci. 15, 1–34.
CCAMLR, 2018a. Fishery Report 2018: Dissostichus eleginoides Heard Island Australian EEZ (Division 58.5.2).
CCAMLR, 2018b. Fishery Report 2018: Dissostichus eleginoides South Georgia (Subarea 48.3).
CCAMLR, 2018c. Fishery Report 2018: Dissostichus eleginoides Crozet Island French EEZ (Subarea 58.6).
CCAMLR, 2018d. Fishery Report 2018: Dissostichus eleginoides Kerguelen Islands French EEZ (Division 58.5.1).
Cisneros-Montemayor, A.M., Cisneros-Mata, M.A., Harper, S., Pauly, D., 2013. Extent and implications of IUU catch in Mexico's marine fisheries. Mar. Policy 39, 283–288.
Clark, J.M., Agnew, D.J., 2010. Estimating the impact of depredation by killer whales and sperm whales on longline fishing for toothfish (Dissostichus eleginoides) around South Georgia. CCAMLR Sci. 17, 163–178.
Collins, M.A., Brickle, P., Brown, J., Belchier, M., 2010. The Patagonian toothfish: biology, ecology and fishery. Advances in Marine Biology. Elsevier, pp. 227–300.
Constable, A.J., De LaMare, W.K., Agnew, D.J., Everson, I., Miller, D., 2000. Managing fisheries to conserve the Antarctic marine ecosystem: practical implementation of the Convention on the Conservation of Antarctic Marine Living Resources (CCAMLR). ICES J. Mar. Sci. 57, 778–791. https://doi.org/10.1006/jmsc.2000.0725.
Cook, R.M., 2013. A fish stock assessment model using survey data when estimates of catch are unreliable. Fish. Res. 143, 1–11.
Evans, S.M., Hunter, J.E., Elizal, Wahju, R.I., 1994. Composition and fate of the catch and bycatch in the Farne Deep (North Sea) Nephrops fishery. ICES J. Mar. Sci. 51, 155–168. https://doi.org/10.1006/jmsc.1994.1017.
Farmer, B., Nowara, G.B., Barnes, T.C., Burch, P., Woodcock, E., Ziegler, P., Welsford, D. C., 2019. Modelling the spatial distribution of Patagonian toothfish (Dissostichus eleginoides) by length and age around Heard Island and McDonald Islands on the Kerguelen Plateau, in: The Kerguelen Plateau: Marine Ecosystem and Fisheries. Proceedings of the Second Symposium. Australian Antarctic Division 219–235.
Fournier, D., Archibald, C.P., 1982. A general theory for analyzing catch at age data. Can. J. Fish. Aquat. Sci. 39, 1195–1207.
Francis, R.I.C.C., 2011. Data weighting in statistical fisheries stock assessment models. Can. J. Fish. Aquat. Sci. 68, 1124–1138. https://doi.org/10.1139/f2011-025.
Francis, R.I.C.C., 2017. Revisiting data weighting in fisheries stock assessment models. Fish. Res. 192, 5–15. https://doi.org/10.1016/j.fishres.2016.06.006.
Gabriel, W.L., Mace, P.M., 1999. A review of biological reference points in the context of the precautionary approach. In: Proceedings of the Fifth National NMFS Stock Assessment Workshop: Providing Scientific Advice to Implement the Precautionary Approach Under the Magnuson-Stevens Fishery Conservation and Management Act. NOAA Tech Memo NMFS-F/SPO-40, pp. 34–45.
Hammond, T.R., Trenkel, V.M., 2005. Censored catch data in fisheries stock assessment. ICES J. Mar. Sci. 62, 1118–1130.
Hanchet, S., Sainsbury, K., Butterworth, D., Darby, C., Bizikov, V., Godø, O.R., Ichii, T., Kock, K.-H., Abellán, L.L., Vacchi, M., 2015. CCAMLR's precautionary approach to management focusing on Ross Sea toothfish fishery. Antarct. Sci. 27, 333.
Hearn, W.S., Polacheck, T., Pollock, K.H., Whitelaw, W., 1999. Estimation of tag reporting rates in agestructured multicomponent fisheries where one component has observers. Can. J. Fish. Aquat. Sci. 56, 1255–1265.
Hilborn, R., Walters, C.J., 1992. Quantitative Fisheries Stock Assessment: Choice, Dynamics and Uncertainty. Chapman and Hall, New York.
Hoenig, J.M., Barrowman, N.J., Hearn, W.S., Pollock, K.H., 1998. Multiyear tagging studies incorporating fishing effort data. Can. J. Fish. Aquat. Sci. 55, 1466–1476.

Hordyk, A.R., Huynh, Q.C., Carruthers, T.R., 2019. Misspecification in stock assessments: common uncertainties and asymmetric risks. Fish Fish. Oxf. (Oxf) 20, 888–902.
Hoyle, S.D., Maunder, M.N., A'mar, Z.T., 2020. Frameworks for the next generation of general stock assessment models: 2019 CAPAM workshop report. New Zeal. Fish. Assess. Rep. 39.
Ichinokawa, M., Okamura, H., Takeuchi, Y., 2014. Data conflict caused by model mis-specification of selectivity in an integrated stock assessment model and its potential effects on stock status estimation. Fish. Res. 158, 147–157.
Magnusson, A., Hilborn, R., 2007. What makes fisheries data informative? Fish Fish. 8, 337–358.
Maunder, M.N., Piner, K.R., 2014. Contemporary fisheries stock assessment: many issues still remain. ICES J. Mar. Sci. 72, 7–18. https://doi.org/10.1093/icesjms/fsu015.
Maunder, M.N., Punt, A.E., 2013. A review of integrated analysis in fisheries stock assessment. Fish. Res. 142, 61–74. https://doi.org/10.1016/j.fishres.2012.07.025.
Mormede, S., Dunn, A., 2013. Quantifying vessel performance in the CCAMLR tagging program: spatially and temporally controlled measures of tag-detection rates. CCAMLR Sci. 20, 73–80.
Omori, K.L., Hoenig, J.M., Luehring, M.A., Baier-Lockhart, K., 2016. Effects of underestimating catch and effort on surplus production models. Fish. Res. 183, 138–145. https://doi.org/10.1016/j.fishres.2016.05.021.
Patterson, K.R., Pitcher, T.J., Stokes, T.K., 1993. A stock collapse in a fluctuating environment: the chub mackerel Scomber japonicus (Houttuyn) in the eastern central Pacific. Fish. Res. 18, 199–218.
Pauly, D., Zeller, D., 2016. Catch reconstructions reveal that global marine fisheries catches are higher than reported and declining. Nat. Commun. 7 https://doi.org/10.1038/ncomms10244.
Pine, W.E., Pollock, K.H., Hightower, J.E., Kwak, T.J., Rice, J.A., 2003. A review of tagging methods for estimating fish population size and components of mortality. Fisheries 28, 10–23.
Pitcher, T.J., Watson, R., Forrest, R., Valtýsson, H.P., Guénette, S., 2002. Estimating illegal and unreported catches from marine ecosystems: a basis for change. Fish Fish. 3, 317–339. https://doi.org/10.1046/j.1467-2979.2002.00093.x.
Plagányi, É.T., Butterworth, D., Burgener, M., 2011. Illegal and unreported fishing on abalone-Quantifying the extent using a fully integrated assessment model. Fish. Res. 107, 221–232. https://doi.org/10.1016/j.fishres.2010.11.005.
Polacheck, T., 2012. Assessment of IUU fishing for southern bluefin tuna. Mar. Policy 36, 1150–1165.
Polacheck, T., Eveson, J.P., Laslett, G.M., Pollock, K.H., Hearn, W.S., 2006. Integrating catch-at-age and multiyear tagging data: a combined Brownie and Petersen estimation approach in a fishery context. Can. J. Fish. Aquat. Sci. 63, 534–548. https://doi.org/10.1139/f05-232.
Pollock, K.H., Hoenig, J.M., Hearn, W.S., Calingaert, B., 2001. Tag reporting rate estimation: 1. An evaluation of the high-reward tagging method. North Am. J. Fish. Manag 21, 521–532.
Pollock, K.H., Hearn, W.S., Polacheck, T., 2002a. A general model for tagging on multiple component fisheries: an integration of age-dependent reporting rates and mortality estimation. Environ. Ecol. Stat. 9, 57–69.
Pollock, K.H., Hoenig, J.M., Hearn, W.S., Calingaert, B., 2002b. Tag reporting rate estimation: 2. Use of high-reward tagging and observers in multiple-component fisheries. North Am. J. Fish. Manag 22, 727–736.
Porch, C.E., Eklund, A.-M., Scott, G.P., 2006. A catch-free stock assessment model with application to goliath grouper (Epinephelus itajara) off southern Florida. Fish. Bull. 104, 89–101.
Punt, A.E., Smith, D.C., Tuck, G.N., Methot, R.D., 2006. Including discard data in fisheries stock assessments: two case studies from south-eastern Australia. Fish. Res. 79, 239–250.
Punt, A.E., Dunn, A., Elvarsson, B.Þ., Hampton, J., Hoyle, S.D., Maunder, M.N., Methot, R.D., Nielsen, A., 2020. Essential features of the next-generation integrated fisheries stock assessment package: a perspective. Fish. Res. 229, 105617.
Ricker, W.E., 1975. Computation and interpretation of biological statistics of fish populations. Bull. Fish. Res. Bd. Can. 191, 1–382.
Roche, C., Guinet, C., Gaseo, N., Duhamel, G., 2007. Marine mammals and demersal longline fishery interactions in crozet and kerguelen exclusive economic zones: an assessment of depredation levels. CCAMLR Sci. 14, 67–82.
Roda, M.A.P., Gilman, E., Huntington, T., Kennelly, S.J., Suuronen, P., Chaloupka, M., Medley, P.A., 2019. A Third Assessment of Global Marine Fisheries Discards. Food and Agriculture Organization of the United Nations.
Rudd, M.B., Branch, T.A., 2017. Does unreported catch lead to overfishing? Fish Fish. 18, 313–323. https://doi.org/10.1111/faf.12181.
Sackett, D.K., Catalano, M., Drymon, M., Powers, S., Albins, M.A., 2018. Estimating exploitation rates in the Alabama red snapper fishery using a high-reward tag–recapture approach. Mar. Coast. Fish. 10, 536–549.
Seber, G.A.F., 1982. The Estimation of Animal Abundance and Related Parameters.
Tixier, P., Gasco, N., Duhamel, G., Viviant, M., Authier, M., Guinet, C., 2010. Interactions of Patagonian toothfish fisheries with killer and sperm whales in the Crozet islands Exclusive Economic Zone: an assessment of depredation levels and insights on possible mitigation strategies. CCAMLR Sci. 17, 179–195.
Tixier, P., Burch, P., Massiot-Granier, F., Ziegler, P., Welsford, D., Lea, M.-A., Hindell, M. A., Guinet, C., Wotherspoon, S., Gasco, N., 2020. Assessing the impact of toothed whale depredation on socio-ecosystems and fishery management in wide-ranging subantarctic fisheries. Rev. Fish Biol. Fish. 30, 203–217.
Tuck, G.N., De La Mare, W.K., Hearn, W.S., Williams, R., Smith, A.D.M., He, X., Constable, A., 2003. An exact time of release and recapture stock assessment model with an application to Macquarie Island Patagonian toothfish (Dissostichus eleginoides). Fish. Res. 63, 179–191.
Van Beveren, E., Duplisea, D., Castonguay, M., Doniol-Valcroze, T., Plourde, S., Cadigan, N., 2017. How catch underreporting can bias stock assessment of and advice for northwest Atlantic mackerel and a possible resolution using censored catch. Fish. Res. 194, 146–154. https://doi.org/10.1016/j.fishres.2017.05.015.
Van Beveren, E., Duplisea, D.E., Marentette, J.R., Smith, A., Castonguay, M., 2020. An example of how catch uncertainty hinders effective stock management and rebuilding. Fish. Res. 224, 105473.
Van der Elst, R., Everett, B., Jiddawi, N., Mwatha, G., Afonso, P.S., Boulle, D., 2005. Fish, fishers and fisheries of the Western Indian Ocean: their diversity and status. A preliminary assessment. Philos. Trans. R. Soc. A Math. Phys. Eng. Sci. 363, 263–284.
Varkey, D.A., Ainsworth, C.H., Pitcher, T.J., Goram, Y., Sumaila, R., 2010. Illegal, unreported and unregulated fisheries catch in Raja Ampat regency, Eastern Indonesia. Mar. Policy 34, 228–236. https://doi.org/10.1016/j.marpol.2009.06.009.
Walters, C., Martell, S.J.D., 2002. Stock assessment needs for sustainable fisheries management. Bull. Mar. Sci. 70, 629–638.
Wang, S.-P., Maunder, M.N., Aires-da-Silva, A., 2009. Implications of model and data assumptions: an illustration including data for the Taiwanese longline fishery into the eastern Pacific Ocean bigeye tuna (Thunnus obesus) stock assessment. Fish. Res. 97, 118–126.
Watson, R., Pauly, D., 2001. Systematic distortions in world fisheries catch trends. Nature 414, 534–536. https://doi.org/10.1038/35107050.
Welsford, D.C., Ziegler, P.E., 2013. Factors that may influence the accuracy of abundance estimates from CCAMLR tag-recapture programs for *Dissostichus* spp. and best practices for addressing bias. CCAMLR Sci. 20, 63–72.
Wilberg, M.J., Thorson, J.T., Linton, B.C., Berkson, J., 2009. Incorporating time-varying catchability into population dynamic stock assessment models. Rev. Fish. Sci. 18, 7–24.
Zeller, D., Darcy, M., Booth, S., Lowe, M.K., Martell, S., 2008. What about recreational catch? Potential impact on stock assessment for Hawaii' s bottomfish fisheries Author' s personal copy 91, 88–97. https://doi.org/10.1016/j.fishres.2007.11.010.
Ziegler, P.E., 2013. Influence of data quality and quantity from a multiyear tagging program on an integrated fish stock assessment. Can. J. Fish. Aquat. Sci. 70, 1031–1045. https://doi.org/10.1139/cjfas-2012-0413.
Ziegler, P.E., Frusher, S.D., Johnson, C.R., 2003. Space–time variation in catchability of southern rock lobster Jasus edwardsii in Tasmania explained by environmental, physiological and density-dependent processes. Fish. Res. 61, 107–123.