DTU Library

Effects of misreporting landings, discards, and Catch Per Unit of Effort index in state-space production models: the case of black hake in northwest Africa

Soto, Maria; Fernandez-Peralta, Lourdes; Grazia Pennino, Maria; Kokkalis, Alexandros; Rey, Javier; Salmeron, Francisca; Liebana, Maria; Meissa, Beyah; Kell, Laurie

Published in:
ICES Journal of Marine Science

Link to article, DOI:
10.1093/icesjms/fsac188

Publication date:
2023

Document Version
Publisher's PDF, also known as Version of record

Citation (APA):
Soto, M., Fernandez-Peralta, L., Grazia Pennino, M., Kokkalis, A., Rey, J., Salmeron, F., Liebana, M., Meissa, B., & Kell, L. (2023). Effects of misreporting landings, discards, and Catch Per Unit of Effort index in state-space production models: the case of black hake in northwest Africa. *ICES Journal of Marine Science*, *80*(10), 2591-2605. Article fsac188. https://doi.org/10.1093/icesjms/fsac188

ICES Journal of Marine Science, 2023, **80**, 2591–2605
DOI: 10.1093/icesjms/fsac188
Advance access publication date: 14 November 2022
Review Article

ICES CIEM International Council for the Exploration of the Sea Conseil International pour l'Exploration de la Mer

Effects of misreporting landings, discards, and Catch Per Unit of Effort index in state-space production models: the case of black hake in northwest Africa

María Soto [1,*], Lourdes Fernández-Peralta [1], Maria Grazia Pennino [2], Alexandros Kokkalis [3], Javier Rey [1], Francisca Salmerón [1], María Liébana [1], Beyah Meissa [4], and Laurie Kell [5]

[1]Instituto Español de Oceanografía (IEO-CSIC), Centro Oceanográfico de Málaga, Puerto Pesquero s/n, Fuengirola, 29640 Málaga, Spain
[2]Instituto Español de Oceanografía (IEO-CSIC), Centro Oceanográfico de Vigo, Subida a Radio Faro 50-52, 36390 Vigo (Pontevedra), Spain
[3]National Institute of Aquatic Resources, Kemitorvet, Building 202, 2800 Kgs. Lyng, Denmark
[4]Institut Mauritanien de Recherche Océanographique et des Pêches (IMROP), BP 22 Cansado, Nouadhibou, Mauritania
[5]Center for Environmental Policy, Imperial College London, South Kensington Campus, London SW7 2AZ, UK

* *Corresponding author: tel: +34952197124;* email: maria.soto@ieo.csic.es.

Recently, various state-space implementations of surplus production models (SPMs) have been developed for data-limited stocks. Often, catches and fishing effort are underestimated and discards are ignored. This results in biased estimates of stock status and reference points. Therefore, we conduct a sensitivity analysis for different under-reporting scenarios (due to non-declared landings, by-catch, and discards) on model estimates and thus advice for the black hake species in northwest Africa. Two modelling frameworks were used, namely a stochastic SPM in continuous time (SPiCT) and Just Another Bayesian Biomass Assessment (JABBA). A common set of diagnostics was developed to allow comparison across modelling frameworks. Scenarios correspond to hypotheses about misreporting and assumptions and priors that were kept consistent. The ratio of current fishing mortality over the fishing pressure that gives the maximum sustainable yield, F/F_{MSY}, is most affected by under-reporting. Results are sensitive to the prior assumed for the initial depletion level, B_0/K, and research is needed. If the misreporting is changing over time, relative quantities (e.g. F/F_{MSY}) and trends are biased, while if misreporting (or at least a part of misreporting) is constant, relative quantities are unbiased. Therefore, the nature of any trend in misreporting should be investigated.

Keywords: data-limited stocks, discards, diagnostics, *Merluccius* spp., state-space models, stock assessment

Introduction

In recent decades, fisheries management policies (DAFF, 2007; MSA, 2007; MFNZ, 2008; CFP, 2013) have paid special attention to improving data-limited stock (DLS) advice (e.g. Froese *et al.*, 2012). This has resulted in the development of a number of computational approaches (Sharma *et al.*, 2021; Cousido-Rocha *et al.*, 2022).

When only total catch and effort data are available for a fishery, surplus production models (SPMs) are the most commonly used tools (Hilborn and Walters, 1992) and have a long history as a method for managing data-limited fish stocks (Punt, 2003). The dynamics of the populations are defined by the relationship between current biomass and previous biomass, a surplus production function (that encompasses the processes of growth, recruitment, and natural mortality), and catch (Polacheck *et al.*, 1993). A common assumption is that a relative abundance index is directly proportional to the biomass.

Although production models may produce parameter estimates based on simpler assumptions than age-structured models, they are still used for the assessments of many stocks, for example, tropical tunas (ICCAT, 2019a), North Sea stocks (ICES, 2021b), Mediterranean stocks (STECF, 2017), and North Atlantic African stocks (FAO, 2020b). Recent progress in the fitting procedures of SPMs includes the use of Bayesian methods to set priors (McAllister and Kirchner, 2001), estimation methods to incorporate both process and observation error using mixed-effects (Thorson and Minto, 2015), and Bayesian state-space models (Meyer and Millar, 1999; Thorson *et al.*, 2014; Pedersen and Berg, 2016; Winker *et al.*, 2018). These developments are particularly useful for improving convergence in data-limited situations. The use of bespoke methods designed for specific fisheries is currently being replaced by the use of publicly-available software packages (Methot and Wetzel, 2013; Pedersen and Berg, 2016; Carruthers and Hordyk, 2018; Winker *et al.*, 2018). These are generally flexible, well documented, tested, and maintained (Dichmont *et al.*, 2021). The Food and Agriculture Organization of the UN (FAO) also performs a systematic assessment of stocks based on global landing records alone (FAO, 2019). Nevertheless, despite the development of statistical methods, software, and the availability of documentation, in some cases, changes from traditional methods have not been made.

Although most stock assessment methods rely on reported landings, the assumption that catches are known without error does not hold for many stocks. One of the main problems with catch and effort data is the existence of illegal, under-reported, and unregulated (IUU) fisheries (Lodge, 2007), especially for economically valuable resources. Furthermore, poorly or unmonitored artisanal fleets, by-catch and discards, particularly for trawl fleets, are important sources of underestimation of catch and effort data. Also, the trend or the degree

Received: 4 November 2021; **Revised:** 2 September 2022; **Accepted:** 19 September 2022

of underestimation can vary with time and, depending on the reasons, increase or decrease, masking or overlapping individual effects of under-reported data. Several studies based on data from on-board observers report that total discards can be substantial and account for as much as 40% of total catch for bottom trawlers, and may also be highly variable (Kelleher, 2005; Pérez-Roda *et al.*, 2019), even within the same fleet. In general, information on discards is uncertain and incomplete in many demersal fleets.

Various studies have evaluated the effects on SPM of not including discards, for example, by assuming bounds for catches in a Bayesian framework (Hammond and Trenkel, 2005), while purely frequentist methods have evaluated the effects of misreported catch on species composition of multispecies fisheries (Soto *et al.*, 2006) and on catch and effort jointly (Omori *et al.*, 2016). Nevertheless, there is no analysis of the effect of misreporting on recently developed Bayesian state-space production models. Therefore, as the demand for assessment of DLS is increasing, it is important to assess the robustness of SPM to under-reported data.

To evaluate the effect of the major uncertainties arising from under-reporting catches, discards, and bycatch on management quantities, we chose two widely used Bayesian state-space models: a Stochastic Surplus Production model in Continuous Time (SPiCT) (https://github.com/DTUAqua/spict) and Just Another Bayesian Biomass Assessment (JABBA) (https://github.com/JABBAmodel). We focused on black hake (*Merluccius* spp.), an economically important stock that is affected by misreporting. Black hake stock is a valuable demersal resource of interest for the European Union (EU) fleets operating in northwest Africa (hereafter NWA), (FAO Area 34), due to its high value and volume of catches (22000 in 2018) (FAO, 2020a).

Recently, a CECAF/FAO Working Group on the assessment of demersal resources in NWA (FAO, 2020b) stressed the need to improve the monitoring of catches, effort, and sizes for black hake, both as a target and as an incidental catch, and to make better estimations of discarded catches to be included in future assessments.

Under-reporting is incorporated into the assessment models by altering total landings and the Catch Per Unit of Effort (CPUE) index. By simulating the ranges of uncertainty and the types of trends, we create different levels of data-limited scenarios in the fishery and evaluate the magnitude of obviating their effect on the reference points used for management.

This study is based on the observers' on-board sampling programme for the Spanish fresh trawler fleet operating in NWA. Data available for black hake make this stock the most representative among those NWA resources that can be evaluated through state-space SPM to provide measures of uncertainty for management quantities. Also, the results will help when conducting assessments for other species.

The study applies a similar methodology to that used in the simulation study on the effect of under-reported catches and effort by Omori *et al.* (2016) for DLS. An aim is to assist in the future development of operating models and the use of life history traits for black hake in a DLS framework (ICES, 2020) and also the need to incorporate discards in FAO assessments.

Within this context, the purposes of this study are: (i) to provide replicable alternative stock assessments based on SPiCT (Pedersen and Berg, 2016) and JABBA (Winker *et al.*, 2018) SPMs for black hakes and other DLS in NWA; (ii) to evaluate the effect of major uncertainties arising from under-reporting catches (by-catch and discards) on the parameters of production models and, hence, on management quantities as well as in model estimates; and (iii) to analyse the relative importance of different under-reporting scenarios in the Bayesian state-space assessment models SPiCT and JABBA.

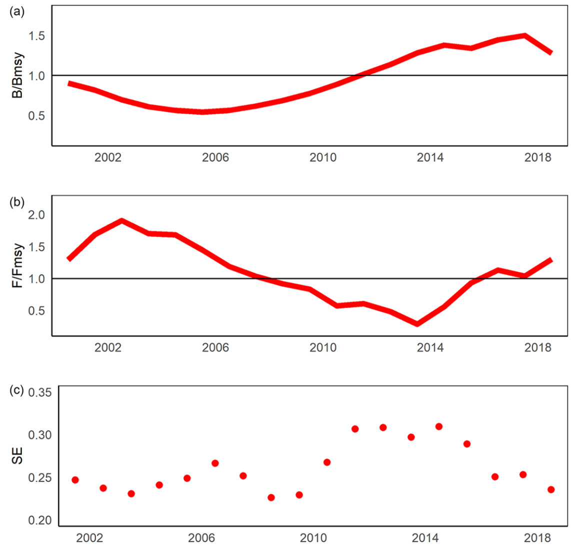

Figure 1. Relative estimates of B_t/B_{MSY} (a) and F_t/F_{MSY} (b) from the Biodyn assessment of black hake in 2019 and standard errors of the standardized CPUE index of the Spanish fresh trawling fleet (c) of black hake.

Material and methods

Black hake fisheries data

The African black hake stock is composed of two sympatric species, *Merluccius polli* (Cadenat, 1950) and *Merluccius senegalensis* (Cadenat, 1950), with fishing ranging from Morocco to Guinea-Bissau. The target fishery is assessed as a single stock, *Merluccius* spp. (FAO, 2020a). The fishery started to extend southward from Saharan waters in the 60s, monitoring began in 1983 and data were collected only from 2000 onward. Fleets operating in the area (EU, coastal states, and other countries) catch black hake as both a target species, there is also discarding of all the fleets operating in the area, the Spanish demersal fleet targeting black hake is the most stable in terms of data continuity and consistency (FAO, 2020b), despite some interruptions in the fishing agreements between EU and non-EU countries. Additionally, this fleet is the only one targeting black hake in the area where observers on board are recording information on catches, discards, and biology. Current assessments of black hake are carried out assuming a Schaefer production model using Biodyn software in discrete time programmed in an EXCEL spreadsheet template (Punt and Hilborn, 1997; Barros, 2012) (Figure 1) fitted assuming an observation error in a frequentist way. *CI* and uncertainty about reference points are not provided. The FAO's current assessment of black hake serves as a baseline assessment for deciding on Total Allowable Catches (TACs) or on fishing effort based on Maximum Sustainable Yield (MSY) reference points, even though it does not follow the minimum scientific criteria required for acceptance (Punt *et al.*, 2020). Estimates of the current Biodyn assessment are shown in Table 1a).

Table 1. (a) Common priors for SPiCT and JABBA assessments and Biodyn estimates used to set priors for the state-space models. (b) Specific initial values, priors and settings for SPiCT and JABBA. (c) Model definitions based on (a) and (b) combining the initial stock status biomass with the use of Standard Errors (SE) of the standardized CPUE index. (d) Notations used in (a), (b), and (c).

(a)	Common priors for SPiCT and JABBA	Biodyn estimates	(b)	Specific priors and settings	SPiCT	JABBA
	Model : Pella-Tomlinson SPiCT `logK ~ c(log(88 954),0.2,1)` `logr ~ c(log(0.59),0.2,1)` B_0`/K ~ logbkfrac = c(log(0.47),0.2)` JABBA `logK = c(log(88 954),0.2)` `logr = c(log(0.59),0.2)` B_0`/K ~ psi.prior = c(0.47, 0.3)`	Model: Schaefer $r = 0.59$ $K = 88954$ $B_{2001}/K = 47\%$ $B_{MSY} = 44\,477$		Initial parameters	`logn = log(2)` `logsdb = 0.2` `logsdc = 0.2` `logbkfrac = log(0.47)`	`init.K = 88 954` `init.r = 0.59` `init.q = 0.1`
				Priors	`logsdi = c(log(0.1699),0.2,1)`	`igamma = c(0.01,0.001)`
				Settings	`dtc = rep(1,length (spict$input$timeC))`	`ni = 5500; nt = 1;` `nb = 500; nc = 2` `proc.dev.all = TRUE` `sigma.add = FALSE` `sigma.est = TRUE` `fixed.obsE = 0.3`

(c)	Model definition	SE CPUE index not used (as Biodyn)	SE CPUE index used
	$B_{MSY}/B_0 = 0.9$ (Biodyn)	<u>Model 1</u> SPiCT `logB = (log(40274.98),0.2,1,2001)` `BmsyB0 = c(log(0.9),0.2,1)` *`SdevfacI = c(1,1,1,1,1,1,1,1,1, 1,1,1,1,1,1,1,1,1)`* JABBA No index SE input	<u>Model 2</u> SPiCT `BmsyB0 = c(log(0.9),0.2,1)` `SdevfacI = c(0.74,0.68,0.65,0.70,0.73,0.77,0.72,0.66,0.68,0.81,0.96,0.98, 0.96,1,0.93,0.81,0.80,0.72)` JABBA `BBmsy = c(0.9,2001,0.2)` `IndexSE = c(0.16,0.15,0.14,0.15,0.16,0.16,0.15,0.14,0.14,0.17,0.21,0.21,0.20, 0.21,0.20,0.17,0.17,0.16)`
	$B_{MSY}/B_0 = 0.5$	<u>Model 3</u> SPiCT `BmsyB0 = c(log(0.5),0.2,1)` `SdevfacI = c(1,1,1,1,1,1,1,1,1,1,1,1, 1,1,1,1,1,1)` JABBA `BBmsy = c(0.5,2001,0.2)` No index SE input	<u>Model 4</u> SPiCT `BmsyB0 = c(log(0.5),0.2,1)` `SdevfacI = c(0.74,0.68,0.65,0.70,0.73,0.77,0.72,0.66,0.68,0.81,0.96, 0.98,0.96,1,0.93,0.81,0.80,0.72)` JABBA `BBmsy = c(0.5,2001,0.2)` `IndexSE = c(0.16,0.15,0.14,0.15,0.16,0.16,0.15,0.14,0.14,0.17,0.21,0.21, 0.20,0.21,0.20,0.17,0.17,0.16)`

Table 1. Continued

(d)	**Description of main symbols used in (a), (b), and (c)**			
	Symbol	Description	Symbol	Description
	bkfrak	ratio of initial biomass to *K*	IndexSE	Vector of the *SD* of the data points of the Observation index
	timeC	Time interval (year) reflecting the beginning of each catch interval	psi.prior	Specify initial depletion prior B_0/K as mean and CV
			igamma	process error variance
	sdb	*SD* of the biomass process observation error	ni	MCMC setting number of iterations
	sdf	*SD* of the fishing mortality process observation error	nt	MCMC setting steps saved
	sdc	*SD* of the catch observation error	nb	MCMC setting burn-in
	sdi	*SD* of the index observation error	nc	MCMC setting number of chains
	dtc	length of each time interval (e.g. 1 for annual)	proc.dev.all	Determines if process error deviation are estimated for all years
	BmsyB0	ratio of initial biomass to B_{MSY}	sigma.add	To estimate additional observation variance
	SdevfacI	vector of factors that are multiplied onto the *SD* of the data points of the corresponding observation index	sigma.est	To estimate additional observation variance
			fixed.obsE	Observation error variance

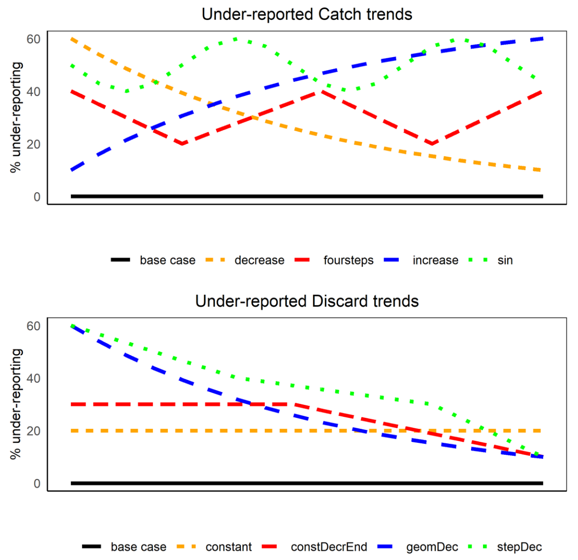

Figure 2. The % of underestimation of total catch and discard of the Spanish trawling fleet for comparison with the current assessments of black hake (*base case* i.e. no underestimation).

For the purpose of this study, a standardized CPUE was developed for the Spanish fresh trawling fleets in Mauritanian waters targeting black hake from 2000 to 2018. It is the same period covered by the above-mentioned Biodyn assessments (FAO, 2020b). Standard errors of the standardized CPUE input are shown in Figure 1. Details of the CPUE standardization are available in the supplementary material. Unlike the Biodyn model, JABBA and SPiCT allow one to incorporate information about the standard errors in the model settings (Table 1c). Total reported catch data for black hake in the NWA (2000–2018) (FAO, 2020b) was used as input for the SPiCT and JABBA production models.

Black hake is a representative example of a fishery with multiple sources of information that results in misreporting of catches and discards. Catch data are occasionally recorded in a heterogeneous spatial-temporal way depending on the country, fleet, or the period of validity of the fishing agreement. This results in a lack of continuity and poor quality of black hake catch and effort data. Also, substantial misreporting target of bycatch catches of non-EU fleets is a recurring concern. Limitations also arise in the sampling at the fish market and by ignoring discards in the assessments (although it is common knowledge that discard numbers are high). Therefore, there is a broad range of possible misreporting scenarios. For example, the underestimation of total catches might have increased due to the increase in power of industrial fleets in recent years, which could generate catches and discards that are ignored in the assessments. Another possibility is that fleets of developing countries are increasingly accessing the fishing grounds. Meanwhile, the Spanish fleet is adopting gradually technical management measures for minimum size. This provoked an increase in discards at the beginning, but fishing strategies are moving to more precautionary levels of discards (for example, fishing in deeper waters). Hence, trajectories of under-reporting discards are assumed to be decreasing (Figure 2).

Production assessment models: priors and settings

Bayesian state-space models relate time series of observations $\{I_t\}$, $\{C_t\}$ to unobserved states, $\{B_t\}$ and $\{F_t\}$ through a stochastic observation model for $\{I_t\}$ given $\{B_t\}$ (Meyer and Millar, 1999). They are able to simultaneously incorporate uncertainty in the biomass dynamics, i.e. process error (Schnute, 1977; Hilborn and Walters, 1992), and the uncertainty in the observations, i.e. observation error (Pella and Tomlinson, 1969; Butterworth and Andrew, 1984; Ludwig and Walters, 1985).

We use the Bayesian state-space models SPiCT and JABBA to assess black hake under different scenarios of under-reporting in the NWA. Both assessment methods implement the Pella-Tomlinson SPM, are distributed as open software written in R (R Development Core Team, 2014), and provide model diagnostics and stock status.

SPiCT is a continuous time model that explicitly models both abundance and fishing dynamics as stochastic processes. The model incorporates observation error not only in the index, σ_I^2, but also in catches, σ_C^2, and process error is associated both with biomass, σ_B^2, and fishing mortality, σ_f^2. Therefore, changes in fishing efficiency will be included in the estimated fishing mortality. SPiCT is implemented using the Template Model Builder (TMB, Kristensen *et al.*, 2015) library in R.

JABBA is a discrete time model that considers the observation error in the index (Winker *et al.*, 2018) σ_ε^2, separated into three additive components for each abundance index i and year y:

$$\sigma^2_{\varepsilon,y,i} = \hat{\sigma}^2_{SE,y,i} + \sigma^2_{fix} + \sigma^2_{est,i},$$

where $\hat{\sigma}^2_{SE,y,i}$ is an external estimable year-to-year observation error, σ^2_{fix}a fix input variance to account for additional sampling errors associated with abundance indices, and $\sigma^2_{est,i}$ an estimable variance from a prior distribution. JABBA is run in JAGS (Plummer, 2003) to estimate Bayesian posterior distributions for all quantities of interest using Monte Carlo Markov Chains (MCMC), and JAGS is executed from the R library R2jags (Su and Yajima, 2012). Process error is associated with the population process, and equations and catch are assumed to have a fixed uncertainty that is not estimated.

Model specifications JABBA and SPiCT are shown in Table 1(a–c).

For SPiCT, a prior for index observation error of the index, σ_I^2 was set at 0.1699 for the mean of the standard errors of the standardized index. The ratios between observation and process errors for the population and the fishery $\alpha = \sigma_I/\sigma_B$ and $\beta = \sigma_C/\sigma_F$ are difficult to estimate simultaneously (Pedersen and Berg, 2016), and observation and process errors were set as default priors. The continuous time equations were discretized using the Euler scheme, with a default time increment of 1/16 year. Also, time for observations of the commercial CPUE was set to the middle of the year, as there is no evidence of seasonality in total catches throughout the year.

For JABBA, after different trials, the prior for process error was set to follow a relative less informative inverse Gamma with scaled parameters (0.01, 0.001) (Winker *et al.*, 2018). We set that the model estimates observation error for all the years, $\hat{\sigma}^2_{SE,y,i}$, i.e. `proc.dev.all = TRUE`, and then estimates additional variance for the observation error $\sigma^2_{est,i}$, i.e. `sigma.add = TRUE` and `sigma.est = TRUE` and set a tentative fix for input variance σ^2_{fix} through `fixed.obsE = 0.3` to account for errors in sampling for the CPUE index.

Model parameters differed between JABBA and SPiCT, therefore, to enable comparisons, when possible common inputs were set for JABBA and SPiCT based on Biodyn assessment estimates (Table 1a) from CECAF (FAO, 2020b). For each model, the input specifications were set following Pedersen and Berg (2016) and Winker *et al.* (2018). The Pella-Tomlinson production function was used rather than the Schaefer used in Biodyn to allow more flexibility in the shape parameter of the production curve. The mean for the prior log-normal distributions for K and r were set to 88952.2 and 0.6, respectively. The prior for the initial depletion level followed a log-normal distribution with mean $= \log(0.47)$ and CV $= 0.3$. The CV of all other parameters in JABBA and SPiCT was set to 0.2.

Model definitions

Model specifications of SPiCT and JABBA have the advantage with respect to Biodyn of setting priors for initial stock status and the ability to provide information about the standard errors of the index as inputs. Initial biomass status estimated from Biodyn was close to the optimum, i.e. $B_{2001}/B_{MSY} = 0.9$ (Figure 1). Nevertheless, as recommended in ICES (2021), a more precautionary value should be considered for initial depletion when it is known that the fishery started many years before data are available (in this case the 80s). For this reason, we pose the case $B_{2001}/B_{MSY} = 0.5$, implemented in JABBA as a prior with a lognormal distribution for B_{2001}/B_{MSY} with CV $= 0.2$.

The Biodyn assessment does not incorporate standard errors for the index. We use this information in JABBA setting `se = data.frame(Yr = timeI, Index = SE)` and in SPiCT as relative uncertainties, i.e. they act as weights in the estimation through `stdevfacI` vector, and are multiplied to the overall index uncertainty that is estimated.

Based on the Biodyn settings and priors combined with the options for initial stock status and index standard errors for JABBA and SPiCT, we obtain four models (Table 1c) for each of which we run sensitivity analysis for underreponting scenarios. The model that replicates the Biodyn assessment is Model 1. The most plausible and precautionary model that incorporates available information is Model 4. Model 2 and Model 3 are intermediate cases between the other two.

Assessment

JABBA and SPiCT are supported and recommended packages when only catch and index data are available (Pedersen and Berg, 2016; Dichcmont *et al.*, 2021). Following Carvalho *et al.* (2021), four properties are used as objective criteria for evaluating the plausibility of the assessment model: (1) model convergence; (2) fit to the data; (3) model consistency; and (4) prediction skill. Here, we focused on the first three to demonstrate the applicability of the JABBA and SPiCT packages. Prediction skill (Kell *et al.*, 2021; Hurtado-Ferro et al. 2015) can be used to choose between models' ability to predict data not used in the fit using the mean absolute scaled error (MASE) (Hyndman and Koelhler, 2006). Regarding the quality of the observable quantities, i.e. the CPUE index, SPiCT, and JABBA are provided with common diagnostic tools and visualization for the input data and CPUE fits, which helps in the evaluation of stock status (Winker *et al.*, 2018) and acceptance of an assessment. The results presented below do not fully describe the complete details of the modelling and prior choices, nor management advice for the NWA black hake stock, but are aimed to illustrate the use of diagnostics and summarize model outputs. Detailed assessment diagnostic plots for the four models are presented in the supplementary material.

Under-reporting scenarios

Under-reporting in total catches and discards may have different patterns over time. We defined four types of trends in under-reporting for each input series in the SPM, $u_{ci}(t)$, (Figure 2 upper panel) for trends in percentages of underreported catches and, $u_{d_i}(t)$, for trends in percentages of under-reported (Figure 2 lower panel).

Percentage of total discards in Spanish trawlers targeting black hake in NWA varies between 20 and 60% (Cervantes *et al.*, 2018; García-Isarch *et al.*, 2020; FAO, 2020a), as recorded by on-board observers. In Mauritania, the discarded black hake by the Spanish fresh trawl fleet has been monitored since 2003 by an on-board sampling programme of observers' surveys (Fernández-Peralta *et al.*, 2010, 2011, 2012; Quintanilla *et al.*, 2012; Rey *et al.*, 2012, 2015, 2016) and the monitoring of the historical evolution of the fishery is described by Ramos and Fernández (1992); Ramos-Martos and Fernández-Peralta (1994, 1995); FAO (2012); Fernández-Peralta *et al.* (2019). In the last decades, black hake discards have progressively decreased, while the size at first capture has increased from 20 to 30 cm in Mauritanian waters (Ramos *et al.*, 1998). To preserve the resource, each country sets a minimum catch size, which influences the amount of hake discarded. This technical measure seems to have resulted in a change of strategy by the fleet, which now fishes deeper to avoid catching juveniles and cause hake discards to diminish. Hence, based on knowledge of the fishery, four trends in under-reported discards were considered: *StepDec*, *geomDec*, *constDecEnd*, and *constant* (Figure 2 lower panel). Under *StepDec*, we assumed three periods of a decrease in discards based on the average effort of the fleet measured in average number of days per trip, 11.5, 7.3, and 6.1, in 2000–2017, in 2008–2014, and in 2015–2018, respectively. In these three periods, the percentage of discards decreased constantly from 60 to 40%, from 40 to 30%, and from 30 to 10%, respectively. In *geomDec*, we considered an exponential decay in the percentage of under-reported discards from 2000 to 2018, from 60 to 10%; in *constDecEnd*, there is a constant initial under-reporting of 30% in the first half of the period and a constant decrease in the second half from 30 to 10% at the end of the period; in *constant*, the under-reported percentage is 30% during the whole period.

Under-reported trends of simulated total catches were defined to cover a range of possible cyclic and monotonic theoretical situations: *sin*, *foursteps*, *decrease*, and *increase* (Figure 2 upper panel). We considered that real reported catches have improved due to the resources invested in data collection, but the global trend in under-reported catches was set as due to many factors (fleets) that made it difficult to assume a predominant trend. To simulate the catch scenarios, we added to the total declared landings, L_t, a percentage given by the under-reporting function:

$$C_{i,t} = \left(1 + u_{c_i}(t)\right) \cdot L_t,\ 1 \leq i \leq 4;\ \ 2000 \leq t \leq 2018.$$

Percentages of under-reported discards, $u_d(t)$, were added as a percentage of the declared hake catches of Spanish Fresh trawl fleet, C_{SP}, and then the numerator in the standardized *CPUE* scenarios were recalculated

$$CPUE_{j,t} = \left(\left(1 + u_{d_j}(t)\right) \cdot C_{t,SP}\right) / f_t, 1 \leq j \leq 4; \quad 2000 \leq t \leq 2018.$$

The current situation (no under-reporting is considered in the assessments) was named the *base case* or Scenario 1 (*S1*), characterized by $u_c(t) = u_d(t) = 0$. The trends of under-reported catches and discards combined generate 16 scenarios "*S2*" to "*S17*" (Table 2) of misreporting to be compared with the *base case*. Note that *decrease* and *geomDec* trends are the same, they only differ in the name to distinguish that one is applied to discards and the other is applied to catches, respectively. Data for scenarios of total catch and CPUE are presented in the supplementary material (Table 1.2).

The scenarios of under-reporting were compared to the base case to investigate how SPMs are affected by uncertainties in total catch, discards, and CPUE, i.e. how important is the observation error in state-space models. Following (Omori *et al.*, 2016), we compared the output of fitting SPiCT and JABBA to each scenario S_i with respect to the *base case* through the percentage difference ($\%difference_Q$) for each parameter, Q,

$$\%difference_Q = 100 \times \frac{(Q_{S_i} - Q_{base})}{Q_{base}}, i = 2, \ldots, 17.$$

Although all model quantities differences are calculated, for the purpose of this study, we focus on the relative differences in relative quantities F/F_{MSY} and B/B_{MSY}.

Results

Assessments

A total of 136 assessments were conducted for the 17 data scenarios, namely, the two state space JABBA and SPiCT models and the four settings established in Model 1–4. Detailed information about these results is available in the online supplementary material. For the purpose of this study, here, we only illustrate some of the most relevant results from the assessments. The assessment criteria of convergence, consistency, and model residuals were achieved for all runs. The Ljun-Box test for autocorrelation of the CPUE residuals was not significant only for SPiCT Model 1, but these do not invalidate the SPiCT results, as only the second lag is significantly high. To illustrate some of the results, Table 3 shows estimates for the *base case* assessments. JABBA and SPiCT *base case* assessments are much more congruent when $B_0/B_{MSY} = 0.5$ and standard errors of the index are used as inputs (Model 4) than when the Biodyn current assessment is basically used (Model 1). Overall, $F/F_{MSY} > 1$, for all models, i.e. the fishing mortality is above the optimal levels. Nevertheless, $B/B_{MSY} > 1$ for all JABBA models and only for SPiCT in Model 4. In general, K and n are lower and r higher for JABBA and SPiCT. The shape parameters decrease from Model 1 to Model 4, and the estimated values are close to Schaefer. MSY is higher for JABBA than SPiCT, except in Model 1. The effect of incorporating the information of standard errors as input is the improvement in relative reference points for both SPiCT and JABBA in all models. Meanwhile, in JABBA, the use of a lower value of 0.5 gives a more pessimistic diagnosis, as it was expected, in SPiCT has the opposite effect, i.e. with lower initial depletion level, lower K and better values of F/F_{MSY}.

Figure 3 shows the prior and posterior distributions of JABBA and SPiCT Model 4. The main differences appear in the prior and posterior of B_0/K. Uncertainty in parameters are difficult to compare between JABBA and SPiCT. The observation error in the index for SPiCT, *sdi*, was parametrized through a log-normal distribution with a mean equal to the average of standard errors of the standardized CPUE, 0.1699, but its estimated posterior distribution differed from the prior (see also supplementary material for more detailed plots). The same occurs to α and β ratios. Regarding JABBA, observation error was estimated using a low informative prior distribution of σ^2_{est} and is very similar between models.

Under-reporting scenarios

Figure 4 shows the significant differences between under-reporting scenarios and the *base case* relative quantities B_{2018}/B_{MSY} and F_{2018}/F_{MSY}. As expected, in the most favourable cases in the under-reported trend of total catches, ($u_d = decrease$), i.e. scenarios S10, S11, S12, and S13, relative reference points resulted in more favourable values of relative biomass and fishing mortality than the *base case*. This means that, independently of the discard under-reporting trend and model considered, the current relative reference points at the end of the period, the improvement in total catch statistics would lead to more optimistic assessments of the stock. Additionally, scenarios with a cyclic under-reporting trend, $u_c = sin$, and decreasing under-reported discards, $u_d =$ *geomdec, constDecrEnd*, i.e. S3 and S4, show the same result in SPiCT and JABBA. The bias generated by the trend of under-reported catches, $u_c = sin$, depends on the SPM model and on the combination with specific under-reported trends in discards, u_d.

Those with *decrease* and *foursteps* trends of under-reported catches, i.e. S6–S13, showed negative bias in B_{2018}/B_{MSY} and positive bias in F_{2018}/F_{MSY}, i.e. a more pessimistic status. Higher biases are produced on F_{2018}/F_{MSY} than on B_{2018}/B_{MSY}. Table 4 shows the average differences in jointly B_{2018}/B_{MSY} and F_{2018}/F_{MSY} by scenario, showing that scenario 14 ($u_c =$ *increasing* trend in underreported catches and $u_d =$ *stepDec* trend in discards) has the more important differences in relative assessment quantities. This is observed in Figure 5. The high levels of non-declared discards of the Spanish fleet at the beginning of the period produce significant differences in catches for the same level of effort and, hence, CPUE trajectories in *base case* and scenario 14 differed notably (shown in the log scale to better visualize the bias). This causes differences in contrast of the data as inputs for the models, mainly at the beginning of the period. This is also observed in the trajectories of total catches at the end of the period, where the increase in declared total catches causes differences with the *base case* to be magnified, producing differences in the contrast of catches. The combination of these two effects is observed in the trajectories of relative quantities such that, at the beginning of the period, B/B_{MSY} is higher for the *base case*, changing along the period to finish below the *base case* B/B_{MSY}. The opposite occurs for the relative F/F_{MSY}.

Figure 6 shows default plots for Kobe plot for the SPiCT *base case* scenario for Model 1 (top panel) and Kobe plot for the JABBA *base case* scenario for Model 4 (bottom panel).

Table 2. Scenarios of under-reporting are generated by the combination of the different trends in under-reporting catches, u_c (*sin*, *foursteps*, *Decrease*, and *Increase*), and under-reporting discards, u_d (*stepDec*, *geomDec*, *constDecend*, *and constant*).

Base case (S1) $u_C = u_D = 0\%$		Catch trend u_C			
		sin	*Foursteps*	*Decrease*	*Increase*
Discard trend u_D	***stepDec***	S2	S6	S10	S14
	geomDec	S3	S7	S11	S15
	constDecEnd	S4	S8	S12	S16
	constant	S5	S9	S13	S17

Table 3. Posterior estimates and *CI* for SPiCT and JABBA assessment of the *base case* (no under-reporting) for Model 1: $B_0/B_{MSY} = 0.9$ without vectors of CPUE standard errors as input; Model 2: $B_0/B_{MSY} = 0.9$ with vectors of CPUE standard errors as input; Model 3: $B_0/B_{MSY} = 0.5$ without vectors of CPUE standard errors as input; and Model 4: $B_0/B_{MSY} = 0.5$ with vectors of CPUE standard errors as input.

		JABBA	SPiCT
Model 1	K	87 670.66 (66844.9–123289.09)	101 415.61 (75880.58–135543.58)
	r	0.63 (0.46–0.85)	0.6 (0.45–0.8)
	N	1.87 (1.33–2.6)	2.17 (1.58–2.98)
	B_{MSY}	42 681.5 (30699.28–63902.32)	52 194.44 (36581.5–74470.95)
	F_{MSY}	0.33 (0.24–0.47)	0.28 (0.21–0.36)
	MSY	14 246.43 (11955.38–18178.77)	14 451.93 (12578.05–16604.98)
	B/B_{MSY}	1.07 (0.75–1.58)	0.93 (0.68–1.26)
	F/F_{MSY}	1.44 (0.82–2.24)	1.57 (0.97–2.54)
Model 2	K	91 432.7 (68067.85–123563.52)	98 930.7 (70181.7–139456.34)
	r	0.62 (0.46–0.85)	0.61 (0.45–0.81)
	N	1.86 (1.31–2.66)	2.11 (1.48–3.01)
	B_{MSY}	44 403.54 (30633.01–63402.75)	50 321.77 (32456.4–78020.98)
	F_{MSY}	0.33 (0.24–0.48)	0.29 (0.2–0.4)
	MSY	14 673.21 (11934.89–19170.53)	14 361.23 (12214.12–16885.76)
	B/B_{MSY}	1.14 (0.73–1.69)	0.98 (0.69–1.39)
	F/F_{MSY}	1.31 (0.73–2.32)	1.49 (0.92–2.43)
Model 3	K	88 404.41 (67472.31–116 164)	98 356.94 (71406.85–135478.42)
	r	0.62 (0.46–0.86)	0.58 (0.44–0.77)
	N	1.86 (1.33–2.54)	2 (1.41–2.83)
	B_{MSY}	42 752.94 (30969.92–60011.25)	49 025.57 (35624.06–73672.82)
	F_{MSY}	0.34 (0.25–0.47)	0.29 (0.21–0.4)
	MSY	14 350.12 (11191.57–18150.73)	14 178.63 (12212.54–16461.25)
	B/B_{MSY}	1.04 (0.7–1.6)	0.94 (0.65–1.37)
	F/F_{MSY}	1.47 (0.82–2.4)	1.56 (0.94–2.61)
Model 4	K	87 019.98 (66075.48–119638.96)	92 937.01 (68471.03–126145.15)
	r	0.63 (0.46–0.87)	0.59 (0.44–0.79)
	N	1.82 (1.27–2.58)	1.92 (1.35–2.73)
	B_{MSY}	41 832.49 (29823.48–61318.77)	45 420.04 (30509.11–67618.49)
	F_{MSY}	0.34 (0.24–0.49)	0.31 (0.22–0.43)
	MSY	14 485.58 (11872.35–18788.52)	13 903.97 (11922.8–16214.34)
	B/B_{MSY}	1.09 (0.69–1.62)	1.01 (0.72–1.5)
	F/F_{MSY}	1.38 (0.78–2.47)	1.45 (0.87–2.4)

This illustrates particular differences between JABBA and SPiCT. For MODEL 1, SPiCT showed the better fits to data (is the only case where there is not a lag 2 AC test of the index) (see supplementary material for details). For this model, in SPiCT the Kobe plot showed the current situation of the *base case* in the red quadrant and JABBA in the orange quadrant, i.e. JABBA and SPiCT estimates showed the major differences for Model 1. On the contrary, Model 4 showed the most congruent results between JABBA and SPiCT with the current status of the fishery and the stock in a risk situation (orange quadrant) for JABBA and SPiCT. Figure 7 shows trajectories of B_{2018}/B_{MSY} and F_{2018}/F_{MSY} for all scenarios for Model 4. General patterns of series are similar, but the year in which there is an inflexion point from which stock begins a recovery period varies among scenarios and JABBA and SPiCT, as well as the sign of the bias in the series.

Discussion

This study evaluated the effect of ignoring discards in the CPUE index and under-reporting of landings in the assessments of the Bayesian state-space SPMs SPiCT and JABBA. As a first in implementing a precautionary approach (Garcia, 1996; Fischer, 2020) we analysed uncertainty on the model parameters, the reference points and the deviations both in the process and the observation error. The assessment models are open-source tools implemented as a flexible assessment framework that we used not only to provide an updated assessment, but also to identify where and how research needs to be strengthened to improve the quality of the assessment for DLS. We applied these models to the black hake fishery in NWA, an example of a DLS affected by under-reporting, in discards, declared catches, and CPUE. In addition, data are

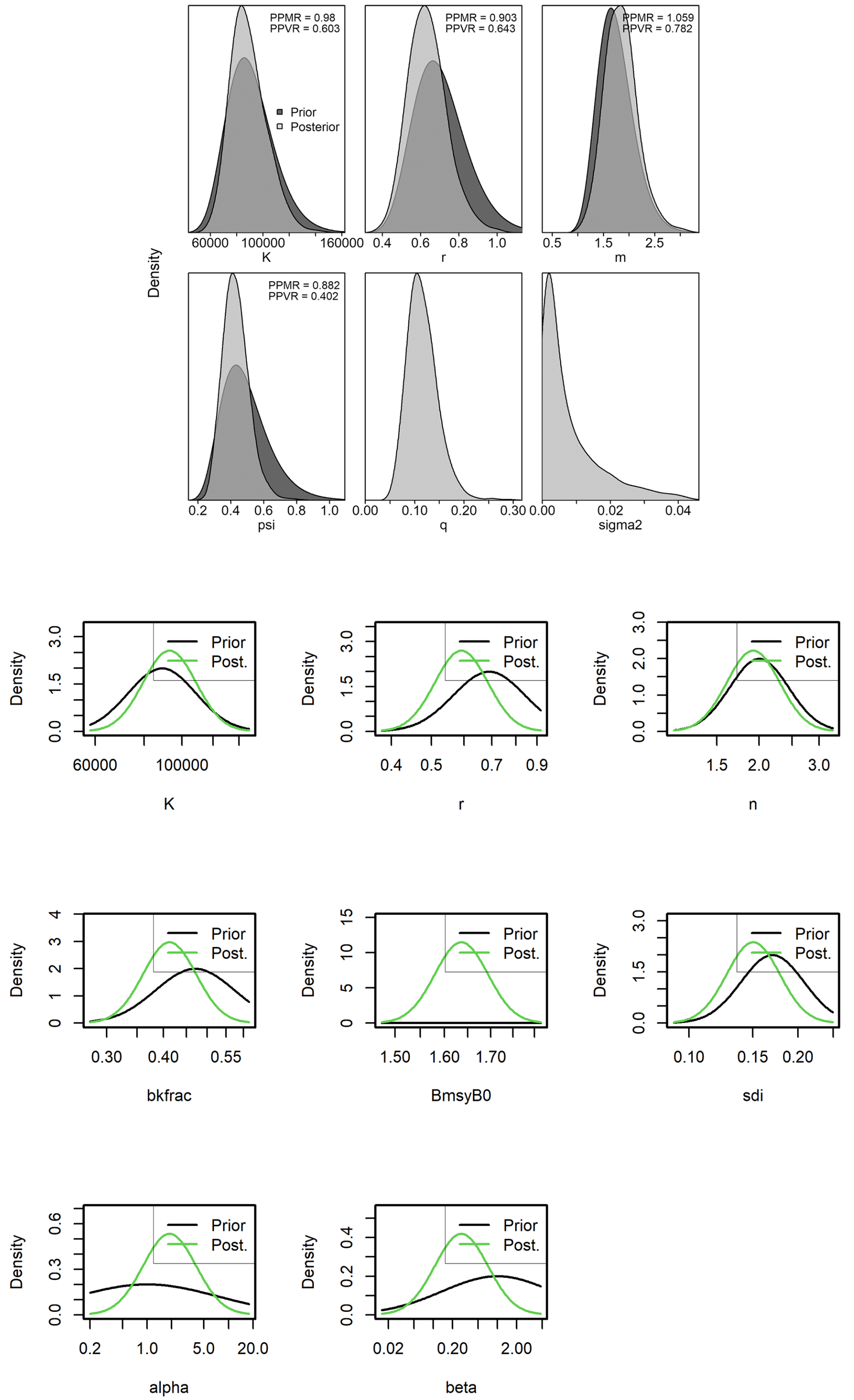

Figure 3. Prior and posterior distributions of parameters estimated in Model 4 for JABBA (top panel) and SPiCT (bottom panel).

uninformative and only short series of data with contrast are available.

The next generation of stock assessment models should routinely provide a set of diagnostics (e.g. Carvalho *et al.*, 2021). We therefore considered criteria to check convergence, goodness of fit, and model consistency in the assessments, which are necessary to make progress in the assessments of black hakes and other DLS. Both models equally identify data issues and indicate that SPiCT assessments are affected by a slight autocorrelation in the index that should be investigated

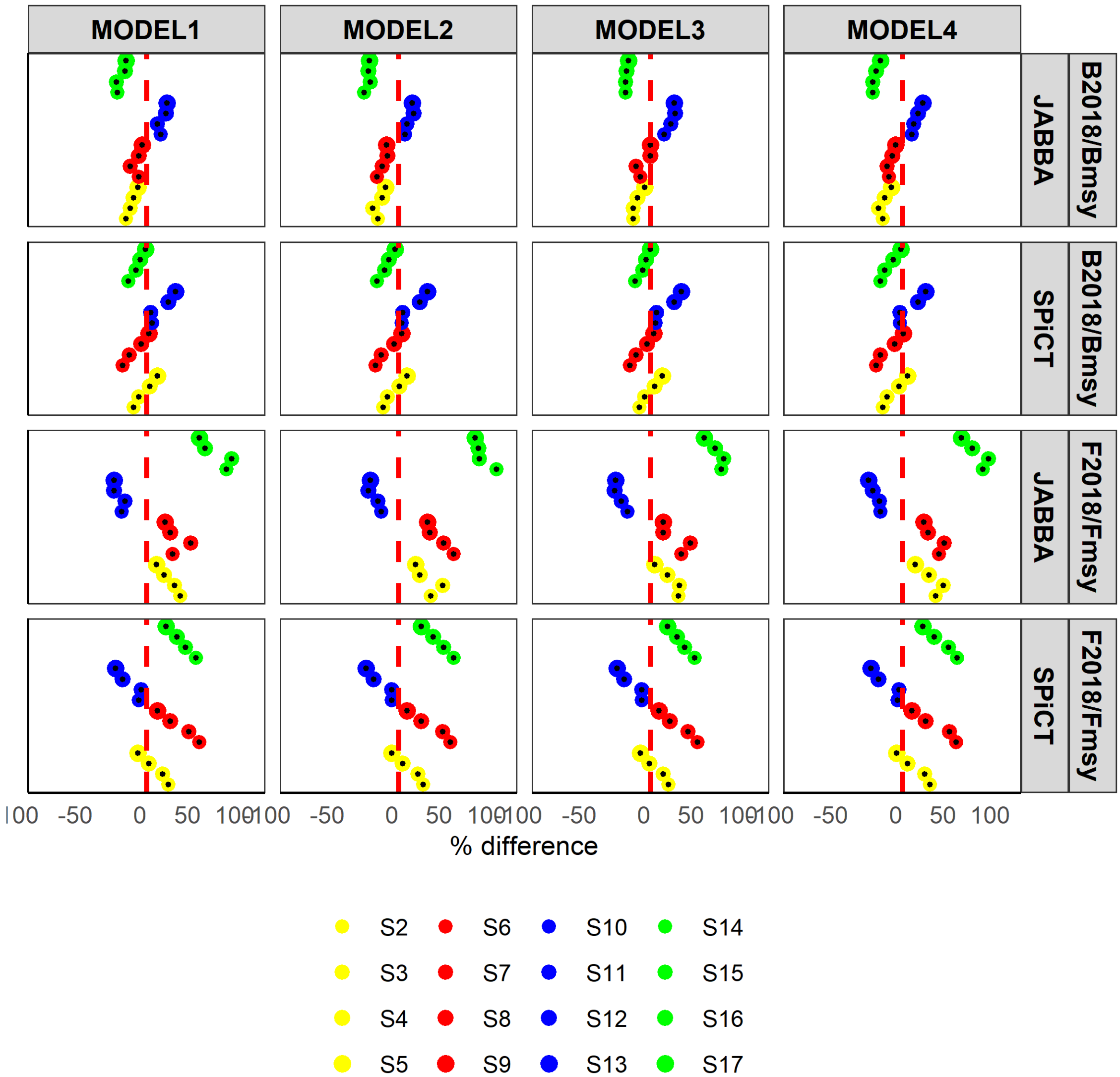

Figure 4. Production model percentage differences in estimates for F/F_{MSY} and B/B_{MSY} between scenarios and the *base case* by model. The dashed red vertical line represents the *base case* reference, and the points at the left and right of the vertical line indicate the sign of the bias in estimates.

Table 4. Average differences between under-reporting scenarios and base case assessments of $F_{current}/F_{MSY}$ or $B_{current}/B_{MSY}$ estimates of JABBA and SPiCT.

Scenario	JABBA	SPiCT
2	5.75	4
3	6.875	3.75
4	2.75	1.875
5	0.75	0.75
6	11.25	13.125
7	12.75	11.75
8	7.625	8
9	6.875	5.5
10	−4.875	−1.75
11	−4.875	−1.25
12	−6	−1.75
13	−5.75	−1.75
14	**24.625**	**14.75**
15	24.75	13.375
16	19.875	11
17	17.375	8.625

Major differences are emboldened and obtained for S14 (*increase* u_c and *stepDec* u_d).

(see supplementary material). The JABBA runs all passed the residual tests. This routine provides a sufficient scientific basis and can be easily implemented to support management advice for black hakes and other stocks in the Eastern Central Atlantic area. Both models present signals of overfishing ($F/F_{MSY} > 1$), being this parameter the most affected by under-reporting.

The definition of the *base case* is a model based on the use of priors (Table 1). The use of priors leads to more robust estimation, but reduces the uncertainty of the estimates (Pedersen and Berg, 2016). Of course, a poor choice of priors that are far from the true values can lead to biased estimates. Recent evolutions in Bayesian computation and software allow for new developments on how to allocate priors in fitting procedures. In particular, state-space models deal with how to combine the uncertainty in the population dynamics with the uncertainty in observed data by using Bayesian techniques to compute posterior distributions of parameters (Meyer and Millar, 1999), and also frequentist methods. SPiCT assigns weights to the standard errors for the index and catches according to the observed variances in the CPUE and catch each

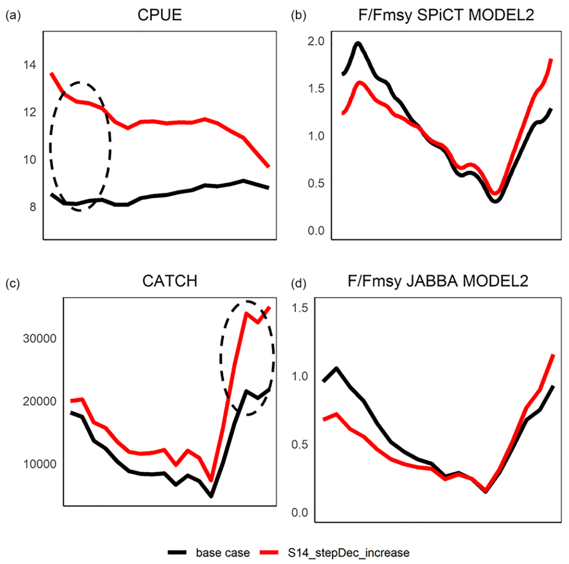

Figure 5. CPUE index (a), catch (c), and estimated relative fishing mortality in SPiCT (b) and JABBA (d) for the original data (*base case*, black line) and *increase* catch together with decreasing step-by-step discards (*stepDec*) of the Spanish fresh trawling fleet scenario (S14) (red line) that yielded the largest % differences in the average B/B_{MSY} and F/F_{MSY} time series. Dashed ovals indicate where the contrast is lowered by under-reporting in discards and total catches.

year. JABBA decomposes the total observation variance of the index allowing for the incorporation of changes in catchability that implicitly informs about the process error in the biomass dynamics. JABBA allows adding the standard errors of the CPUE as inputs. Information about the initial depletion level and status of the stock is crucial to determining the results of the assessments. Posterior distributions for B_0/K differed from priors more than other parameter specifications, both in JABBA and SPiCT.

Starting from the current Biodyn assessment (the most data-limited situation), we set a group of different levels of data-limited situations, combining choices of initial relative stock status with choices about the use of standard errors of the standardized CPUE index, leading to the more plausible situation given by Model 4. This model represents a lower degree of data-limited situation, using information from the index and limiting to a more precautionary level the initial biomass at the beginning of the assessment period. As expected, Model 4 showed the more congruent advice between JABBA and SPiCT. Since both models offer different choices of priors and settings, they are not easy to compare. Nevertheless, some results went beyond control inputs, and we obtained similar conclusions.

Regarding model parameter estimates, the intrinsic population growth rate, r, has a different meaning between the Biodyn, JABBA, and SPiCT approaches. SPiCT is a continuous-time model that is discretized using an Euler scheme. Both these modelling and optimization approaches mean that r is not comparable, and so using the same priors or comparing estimates across models is difficult. A high carrying capacity could be expected in the black hake stock given the resilience of these species, which have endured high fishing pressure with high yields for decades (Pitcher and Alheit, 1995; Rey, 2016). Nevertheless, this is only observed in general for SPiCT, not for JABBA.

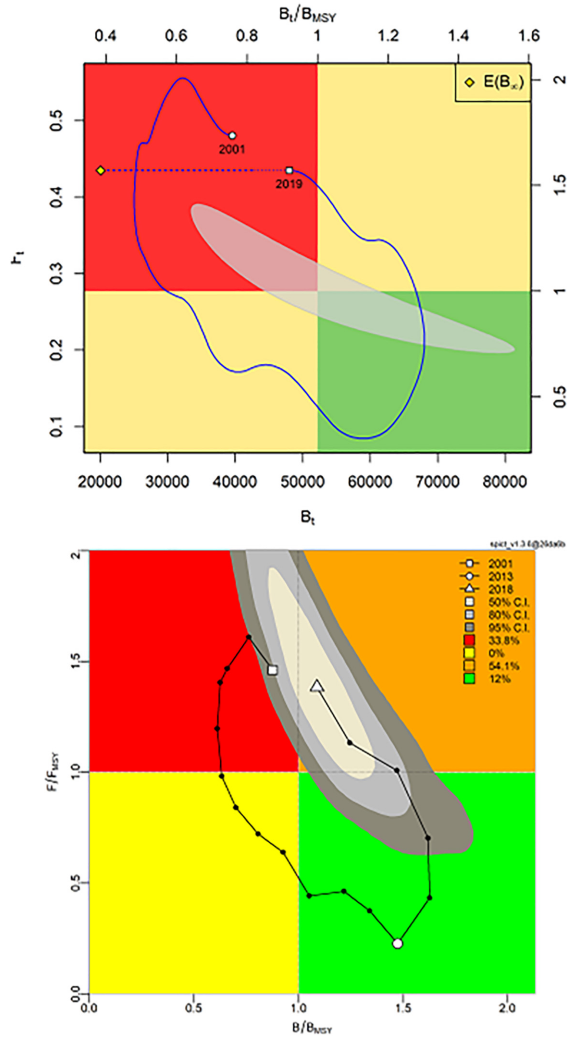

Figure 6: The Kobe plot for the best fit of the SPiCT *base case* scenario for Model 1 (top panel) and the Kobe plot for the JABBA *base case* scenario more similar to SPiCT advice for Model 4 (bottom panel). Plots show the estimated trajectories (2001–2018) of B/B_{MSY} and F/F_{MSY}. Different grey shaded areas in JABBA denote the 50, 80, and 95% credibility intervals for the terminal assessment year. The probability of terminal year points failing within each quadrant is indicated in the figure legend in JABBA. The expected situation if the current levels of fishing mortality are maintained is shown in the yellow point in the SPiCT plot.

As it is recommended by ICES (2021a), this study corroborates that reducing the initial depletion level as well as the relative initial stock status, B_{2001}/B_{MSY}, made assessments of JABBA and SPiCT more congruent. The shape parameter estimated by JABBA is systematically lower than SPiCT, which should be investigated. A balance between model convergence and flexibility to estimate r, K, and n is the key to determining the height, range, and symmetry of the surplus production curve and, hence, the reference points F_{MSY} and B_{MSY} (Pedersen and Berg, 2016). One step ahead, residuals quantify the prediction skill of the SPiCT assessments. We have not implemented a quantification of the predictive skill using MASE, however, as suggested in Kell *et al*. (2021). Also, Winker *et al*. (2018) pinpointed the overoptimistic stock status estimations using JABBA and incorporated a generic hockey-stick function to prevent surplus production per unit of biomass approaching infinity at very low abundance. Nevertheless, we

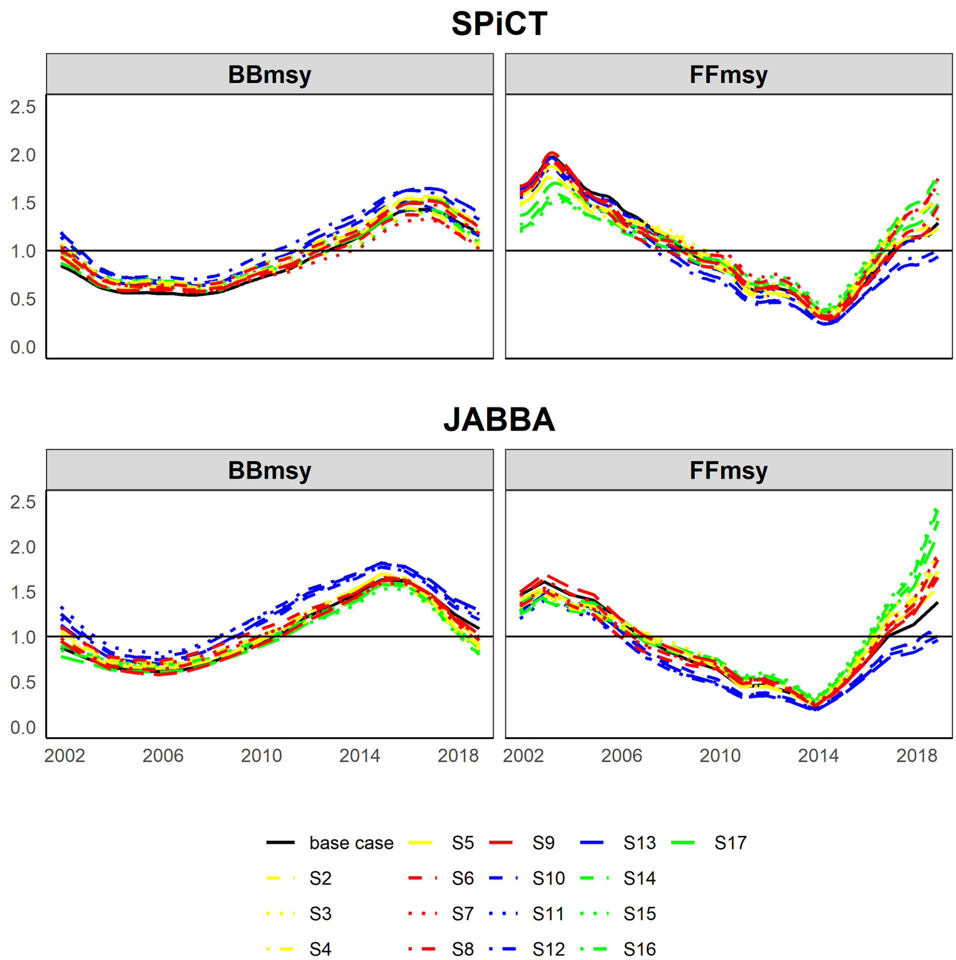

Figure 7. Plot of series B/B_{MSY} and F/F_{MSY} by scenarios and model for SPiCT (top panels) and JABBA (lower panels) for Model 4.

did not use the hockey-stick model, as the shape parameter is close to 2 for all models.

Once the *base case* is defined in our study, scenarios of under-reporting in total catches and discards were incorporated into the SPM, as suggested by Omori *et al.* (2016), with the aim of evaluating observation error to assess systematic bias on estimates of Bayesian state-space models. The difference in this study is that we do not introduce underestimation of effort but of discards. This produced changes in the numerator of the calculation of the CPUE, i.e. at the same levels of effort, catches are set higher due to under-reporting. This changes the patterns of the CPUE at the beginning of the period, showing an overall decreasing pattern (Figure 1) or, at least, a less pronounced increasing trend in the current CPUE. As a consequence, the combination of underestimation trends of total catches and discards affects the contrast of the data (Carvalho et al. 2017) at the beginning of the period, mainly by the higher levels of discards, and at the end of the period due to the higher catches. The result is that the under-reporting effect propagates throughout all the periods, following different patterns depending on the scenario. Overall, the different parametrization affects specific model parameter estimates more (Table 3) and differences in contrast generated by the scenarios drive the advice in JABBA and SPiCT (Figure 4). This means that differences in contrast of the data introduced by under-reporting scenarios create differences in bias of the estimates, independently of the common parametrization used in JABBA and SPiCT.

The impact of systematic trends of under-reporting in catch and discards generates bias in advice with the associated risk. If the trend in under-reported catches is the most optimistic and less realistic case, i.e. the *decrease* trend, and if u_c = *decrease*, the bias in estimated current biomass is positive, i.e. the real biomass would be at higher levels than the *base case*. Nevertheless, this is an improbable situation, and it is more likely to assume that u_c is not decreasing due to a variety of factors. In that framework, a negative systematic bias in current biomass and a more optimistic advice is happening. This means that while we think that the fishery and the stock are near the *MSY* values, the true situation is worst. This conclusion is evident for JABBA. Although in SPiCT most of the scenarios are in the same direction, there are some exceptions. From these results, we conclude that JABBA captures the bias in the data slightly better than SPiCT.

The lack of historical data for many DLS conditions is the starting point of the assessments and highlights the importance of investigating initial values for depletion and proxies of B_0/K in state-space models when short time series are available (around 20 years). These studies should be combined with data on the life history characteristics of the stock (Fischer, 2020) and the quality of past information (through the availability and quality of an abundance index).

In the absence of fishery-independent resource surveys for black hake in NWA, CPUE is currently the best indicator of biomass trends (Cooke and Beddington, 1984; Punt, 2003). For many DLS, particularly in the Central Eastern Atlantic, exploring information obtained through VMS and scientific observation in commercial surveys and also further studies on the hydrological conditions, i.e. upwelling index, (Meiners, 2007; Meiners *et al.*, 2010) are the unique opportunities for providing coherent measures of fishing effort and obtaining proper CPUE indices.

Many of the conclusions of this study are likely to be relevant for other DLS assessed through production models and, particularly, for Eastern Central Atlantic stocks. However, the type of statistical diagnostic and plots used to evaluate whether a proposed assessment is the "best available science" (Punt *et al.*, 2020) differ between jurisdictions. Some stocks are not so data-limited (ICES, 2012), and assessments conducted by CECAF could be improved using modern software and existing data. State-space models fit well and may detect anomalies in the indices. As above mentioned, the way to improve data quality is by using VMS to standardized CPUEs, investigating priors for B_0/K, and collecting historical data on the fisheries and any other relevant information, such as local surveys in cooperation with third countries. Systematic catch under-reporting will affect historical and current estimates and reference points for catch, biomass, and harvest rate. If misreporting has been consistent over time, i.e. a constant percentage, then although the estimate of scale will be biased, relative estimates such as F/F_{MSY} and B/B_{MSY}, both current and historical trends, will potentially be unbiased. However, it is unlikely that misreporting will be consistent over time as it will depend upon monitoring, control, and compliance. For example, the initial development of a fishery when a stock was lightly exploited and a comprehensive management framework lacking misreporting could result in an underestimation of virgin biomass and productivity at high stock sizes (e.g. the increase catch scenario). While misreported in recent years as management measures are implemented as stock becomes deleted (e.g. the decline catch scenario), it will potentially impact estimates of F_{MSY} and B_{MSY}. The precautionary approach requires the impact of uncertainty on estimates of stock status to be taken into account when providing management advice. Catch misreporting is a problem in many fisheries worldwide and understanding the impact is necessary to provide robust advice.

This work contributes to increasing transparency, reproducibility, and scientific reliability, while the external peer-review enhances the consistency of the assessments. The majority of SPMs implemented in Regional Fisheries Management Organizations (RFMO) for stock assessments are based on third-party software (ASPIC; Prager, 1994; BSP2; McAllister, 2014; Biodyn, Barros, 2012). However, changes and further developments to such programmes rely on a few developers, and the record of the issues addressed may be unclear (Winker *et al.*, 2018). Open-source alternatives provide fisheries scientists with tools to improve and standardize assessment procedures, democratizing modelling approaches across nations.

Acknowledgements

We would like to thank the ICES WKLIFE (Workshop on Development of Quantitative Assessment Methodologies based on LIFE-history traits, exploitation characteristics, and other relevant parameters for data-limited stocks) workshop series, which rooted this initiative. Also, the developers of SPiCT and JABBA tutorials crucially contributed to this work by disseminating Bayesian state-space models throughout the scientific community. The comments of the two anonymous reviewers of this document have been enormously constructive for the improvement of this work. Our gratitude also goes to shipowners and fishermen of Spanish fishing trawl vessels operating in Moroccan, Mauritanian, and Senegal waters, and Spanish observers onboard whose work contributed to discards knowledge. Also, Lucía Rueda has contributed with comments to the JABBA settings. MGP also thanks the IMPRESS (RTI2018-099868-B-I00) project, ERDF, Ministry of Science, Innovation and Universities—State Research Agency.

Supplementary data

Supplementary material is available at the *ICESJMS* online version of the manuscript.

Conflict of interest

The authors declare no conflict of interest.

Funding

This work has been carried out in the framework of the structural project of the Instituto Español de Oceanografía, named EPAFRIK-BADE. This project has been co-funded by the EU through the European Maritime and Fisheries Fund (EMFF) within the National Program for the collection, management, and use of data from the fishing sector and support for scientific advice in relation to the Common Fisheries Policy.

Data availability statement

The data underlying this article are available in its online data file "DATA_SOTO.RData" as well as the supplementary material in its online version material.

References

Barros, P., . 2012. Appendix II-Part 1: Biomass dynamic model with environmental effects. User instructions: 251-262.. . FAO. Report of the FAO/CECAF Working Group on the assessment of demersal resources - Subgroup north. Banjul, Gambia, 6–14 November 2007. CECAF/ECAF SERIES 10/71.

Butterworth, D. S., and Andrew, P. A. 1984. Dynamic catch–effort models for the hake stocks in ICSEAF divisions 1.3 to 2.2. International Commission for the Southeast Atlantic Fisheries, 11: 29–58.

Cadenat, J. 1950. Note sur les merlus de la côte occidentale d'Afrique. In: Congr. Pêche Pêcher. Univ. franç. d'outre Mer Inst. Colon., Marseille. pp. 128–130.

Carruthers, T. R., and Hordyk, A. R. 2018. The data-limited methods toolkit (DLMtool): an R package for informing management of data-limited populations. Methods in Ecology and Evolution, 9: 2388–2395.

Carvalho, F., Winker, H., Courtney, D., Kapur, M., Kell, L., Cardinale, M., Schirripa, M. *et al.*. 2021. A cookbook for using model diagnostics in integrated stock assessments. Fisheries Research, 240: 105959.

Carvalho, F., Punt, A. E., Chang, Y.-J., Maunder, M. N., and Piner, K. R. 2017. Can diagnostic tests help identify model misspecification in integrated stock assessments? Fisheries Research, 192: 28–40.

A. Cervantes, A. Faraj, and L. Enthoven (eds.) 2018. Rapport de la Réunion Scientifique Conjointe annuelle relative à l'Accord de pêche signé entre le Royaume du Maroc et l'Union européenne. Madrid, Espagne, 24–26 September 2018. Rapports des Comités Scientifiques Conjoints, Bruxelles. 88pp.

CFP. 2013. Regulation (EU) No 1380/2013 of the European Parliament and of the Council of 11 December 2013 on the Common Fisheries Policy, amending Council Regulations (EC) No 1954/2003 and (EC) No 1224/2009 and repealing Council Regulations (EC) No 2371/2002 and (EC) No 639/2004 and Council Decision 2004/585/EC. Off. J. Eur. Union *L* . 354: 22–61.

Cooke, J. G., and Beddington, J. R. 1984. The relationship between catch rates and abundance in fisheries. Mathematical Medicine and Biology, 1: 391–405.

Cousido-Rocha, M., Cerviño, S., Fernández, A., Gil, J., Gonzalez Herraiz, I., Rincón Hidalgo, M., Ramos, F. *et al.* 2022. Applying length-based assessment methods to fishery resources in the Bay of Biscay and Iberian Coast ecoregion: Stock status and parameter sensitivity. Fisheries Research, 248: 106197.10.1016/j.fishres.2021.106197.

DAFF. 2007. Commonwealth Fisheries Harvest Strategy: Policy and Guidelines. Australian Government, Department of Agriculture, Fisheries and Forestry. http://www.agriculture.gov.au/SiteCollectionDocuments/fisheries/domestic/hsp.pdf.

Dichmont, C. M., Deng, R. A., Dowling, N., and Punt, A. E. 2021. Collating stock assessment packages to improve stock assessments. Fisheries Research, 236: 105844.

FAO. 2012. Report of the FAO/CECAF Working Group on the Assessment of Demersal Resources—Subgroup North. Banjul. CECAF/ECAF Series No. 10/71. Rome, 302pp. https://www.fao.org/publications/card/en/c/CA1003B/ (last accessed 6–14 November 2007).

FAO. 2019. Report of the Expert Consultation Workshop on the Development of Methodologies for the Global Assessment of Fish Stock Status, 4–6 February 2019. FAO Fisheries and Aquaculture Report No. 1262; License CC-BY-NC-SA 3.0 IGO. FAO, Rome. http://www.fao.org/3/ca4355en/ca4355en.pdf.

FAO. 2020a. FAO Yearbook. Fishery and Aquaculture Statistics 2018. Rome. https://doi.org/10.4060/cb1213t. 110pp.

FAO. 2020b. Report of the FAO/CECAF Working Group on the Assessment of Demersal Resources—Subgroup North Nouakchott, Mauritania, 2–10 December 2019 /Rapport du Groupe de travail FAO/COPACE sur l'évaluation des ressources démersales—Sous-groupe Nord Nouakchott, Mauritanie, 2–10 December 2019. https://doi.org/10.4060/cb1539b.

Fernández-Peralta, L., Salmerón, F., Rey, J., Gómez, M. J., García, R., and Macías, D. 2010. Preliminary data on the ovarian histological structures observed in black hakes (*M.polli* and *M. senegalensis*) off Mauritania. *In* Proceedings Workshop on Gonadal Histology of Fishes. Ed. by Wyanski D.and and Brown-Peterson N.. pp. Puerto de Santa María, Cádiz. pp. 191–194.

Fernández-Peralta, L., Salmerón, F., Rey, J., Puerto, M. A., and García-Cancela, R. 2011. Reproductive biology of black hakes (*Merluccius polli* and *M . senegalensis*) off Mauritania. Ciencias Marinas, 37: 527–546.

Fernández-Peralta, L., Salmerón, F., Rey, J., Puerto, M. A., García-Cancela, R., Presas-Navarro, C., and Quintanilla, L. 2012. Distribución de las merluzas negras (*Merluccius senegalensis* y *Merluccius polli*) en aguas de Mauritania. *In* Book of Abstracts of III International Symposium in Marine Science, Cádiz (Spain), 24–27 January 2012, 55pp. ISBN: 978-84-695- 1394-1.

Fernandez Peralta, L., Bouzouma, M., Balguerias, E., Braham, C., Brahim, K., Corten, A., Dia, M. *et al.*. 2019. Rapport de la Réunion annuelle du Comité Scientifique Conjoint relatif à l'Accord de pêche signé entre la République islamique de Mauritanie et l'Union européenne. Santa Cruz de Tenerife, Espagne, Rapports des Comités Scientifiques Conjoints. Bruxelles, 66pp. chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://oceans-and-fisheries.ec.europa.eu/system/files/2021-04/report-jsc-mauritania-2017-10_fr.pdf.

Fischer, S. H., De Oliveira, J. A. A., and Kell, L. T. 2020. Linking the performance of a data-limited empirical catch rule to life-history traits. ICES Journal of Marine Science, 77: 1914–1926.

Froese, R., Zeller, D., Kleisner, K., and Pauly, D. 2012. What catch data can tell us about the status of global fisheries. Marine Biology, 159: 1283–1292

García-Isarch, E., Clark, J. M., Fernández-Peralta, L., González-Lorenzo, J. G., Duque-Nogal, V., Corten, Ad, Rey, J. *et al.*. 2020. Study on Improvement for the Analysis and Exploitation of Observer Reports in EU Fisheries from NW African Waters. Final Report. May 2020, 190pp. Luxembourg: Publications Office of the European Union, 2020, ISBN 978-92-9460-200-8, doi:10.2826/435195.

García, S. 1996. The precautionary approach to fisheries and its implications for fishery research, technology and management: an updated review. FAO Fisheries Technical Paper, pp. 1–76. https://www.fao.org/3/v8045e/V8045E02.htm.

Hammond, T. R., and Trenkel, V. M. 2005. Censored catch data in fisheries stock assessment. ICES Journal of Marine Science, 62: 1118–1130.

Hilborn, R., and Walters, C. J. 1992. Quantitative Fisheries Stock Assessment: Choice, Dynamics, and Uncertainty. Chapman & Hall, New York, NY.

Hurtado-Ferro, F., Szuwalski, C. S., Valero, J. L., Anderson, S. C., Cunningham, C. J., Johnson, K. F., Licandeo, R. *et al.* 2015. Looking in the rear-view mirror: bias and retrospective patterns in integrated, age-structured stock assessment models. ICES Journal of Marine Science, 72: 99–110.

Hyndman, R. J., and Koehler, A. B. 2006. Another look at measures of forecast accuracy. International Journal of Forecasting, 22: 679–688.

ICCAT. 2019. Report of the 2019 ICCAT Yellowfin Tuna data preparatory meeting. https://www.iccat.int/Documents/Meetings/Docs/2019/REPORTS/2019_YFT_DP_ENG.pdf.

ICES. 2012. ICES Implementation of Advice for Data-limited Stocks in 2012 in its 2012 Advice. ICES Document CM 2012 /ACOM: 68. 42pp.

ICES. 2020. Tenth Workshop on the Development of Quantitative Assessment Methodologies based on LIFE-history traits, exploitation characteristics, and other relevant parameters for data-limited stocks. (WKLIFE X). ICES Scientific Reports. 2:98. 72pp. https://doi.org/http://doi.org/10.17895/ices.pub.5985

ICES. 2021a. Benchmark Workshop on the Development of MSY Advice for Category 3 Stocks Using Surplus Production Model in Continuous Time; SPiCT (WKMSYSPiCT). ICES Scientific Reports. 3:20. 317pp. https://doi.org/10.17895/ices.pub.7919.

ICES. 2021b. Benchmark Workshop on North Sea Stocks (WKNSEA). ICES Scientific Reports. 3:25. 756pp. https://doi.org/https://doi.org/10.17895/ices.pub.79229

Kell, L., Shrarma, R., Kitakado, T., Winker, H., Mosqueira, I., Cardinale, M., and Fu, D. 2021. Validation of stock assessment models using prediction skill: is it me or my model talking?. ICES Journal of Marine Science, 78: 2244–2255.

Kelleher, K. 2005. Discards in the world's marine fisheries. An update. FAO Technical Paper No 407, 131. https://doi.org/10.1136/ebmed-2011-0012.

Kristensen, K., Nielsen, A., Berg, C. W., Skaug, H., and Bell, B. 2015. *TMB:* Automatic Differentiation and Laplace Approximation. https://doi.org/10.18637/jss.v070.i05.

Lodge, M. 2007. Managing International Fisheries: Improving Fisheries Governance by Strengthening Regional Fisheries Management Organizations (Briefing Paper). Energy, Environment and Development Programme (Chatham House) 07/01, BP 07/01.

Ludwig, D., and Walters, C. J. 1985. Are age-structured models appropriate for catch–effort data? Canadian Journal of Fisheries and Aquatic Sciences, 42: 1066–1072.

McAllister, M. K., and Kirchner, C. H. 2001. Development of Bayesian stock assessment methods for Namibian orange roughy *Hoplostethus atlanticus*. South African Journal of Marine Science, 23: 241–264.

McAllister, M. K. 2014. A generalized Bayesian surplus production stock assessment software. Collect Vol Sci Pap ICCAT, 70: 1725–1757.

Meiners, C. 2007. Importancia De La Variabilidad ClimáTica En Las PesqueríAs Y BiologíA De La Merluza Europea (*Merlucccius Merluccius* L.) De La Costa Noroccidental Africana. Ph.D thesis, Universitat Politècnica de Catalunya, Catalonia. 206pp.

Meiners, C., Fernández, L., Salmerón, F., and Ramos, A. 2010. Climate variability and fisheries of black hakes (*Merluccius polli* and *M erluccius senegalensis*) in NW Africa: a first approach. Journal of Marine Systems, 80: 243–247.

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

Meyer, R., and Millar, R. B. 1999. Bugs in Bayesian stock assessments. Canadian Journal of Fisheries and Aquatic Sciences, 56: 1078–1087.

MFNZ. 2008. Harvest Strategy Standard for New Zealand Fisheries. Ministry of Fisheries, Wellington.

MSA. 2007. Magnuson-Stevens Fishery Conservation and Management Act, Public Law 94–265. As amended by the Magnuson-Stevens Fishery Conservation and Management Reauthorization Act (P. L. 109–479). http://www.nmfs.noaa.gov/msa2005/docs/MSA_amended_msa%20_20070112_FINAL.pdf.

Omori, K. L., Hoenig, J. M., Luehring, M. A., and Baier-Lockhart, K. 2016. Effects of underestimating catch and effort on surplus production models. Fisheries Research, 183: 138–145.

Pedersen, M. W., and Berg, C. W. 2016. A stochastic surplus production model in continuous time. Fish and Fisheries, 18: 226–243.

Pella, J. J., and Tomlinson, P. K: 1969. A generalized stock production model. Inter-Amarican Tropical Tuna Commission, 13: 421–488.

Pérez-Roda, A., Gilman, E., Huntington, T., Kennelly, S. J., Suuronen, P., Medley, P., Huntington, T. *et al.*. 2019. A third assessment of global marine fisheries discards. *In* FAO Fisheries and Aquaculture Technical Paper No. 633. 78 pp. Licence: CC BY-NC-SA 3.0 IGO. www.fao.org/.

Pitcher, T. J., and Alheit, J. 1995. What makes a hake? A review of the critical biological features that sustain global hake fisheries. *In* Hake: Fisheries, Ecology and Markets, pp. 1–14. Ed. by Alheit J.and and Pitcher T. J.. Springer, Dordrecht. 487pp.

Plummer, M. 2003. JAGS: A Program for Analysis of Bayesian Graphical Models Using Gibbs Sampling. http://www.r-project.org/conferences/DSC-2003/.

Polacheck, T., Hilborn, R., and Punt, A. E. 1993. Fitting surplus production models: comparing methods and measuring uncertainty. Canadian Journal of Fisheries and Aquatic Sciences, 50: 2597–2607.

Prager, M. 1994. A suite of extensions to a nonequilibrium surplus-production model. Fishery Bulletin—National Oceanic and Atmospheric Administration, 92: 374–389.

Punt, A. E. 2003. Extending production models to include process error in the population dynamics. Canadian Journal of Fisheries and Aquatic Sciences, 60: 1217–1228.

Punt, A. E., and Hilborn, R. 1997. Fisheries stock assessment and decision analysis: the Bayesian approach. Reviews in Fish Biology and Fisheries, 7: 35–63.

Punt, A. E., Tuck, G. N., Day, J., Canales, C. M., Cope, J. M., De Moor, C. L., De Oliveira, J. A. A. *et al.*. 2020. When are model-based stock assessments rejected for use in management and what happens then? Fisheries Research, 224: 105465.

Quintanilla, L., Fernández-Peralta, L., Rey, J., García-Cancela, R., Salmerón, F., Puerto, M. A., Presas-Navarro, C., and Saber, S. 2012. Analysis of the Spatio-Temporal Distributions of Black Hakes (*Merluccius polli* and *Merluccius senegalensis*) Abundance in the Mauritanian Coast Using Generalised Additive Models. Foro Iberoamericano de Recursos Marinos y la Acuicultura, Cádiz (España). Libro de Resúmenes del V Foro Iberoamericano de los Recursos Marinos y la Acuicultura, Cádiz. 137pp.

Ramos, A., and Fernández, L. 1992. La pesquería española de arrastre de merluzas negras en aguas Mauritanas: análisis de la serie histórica de datos. Inf Téc Inst Esp Oceanogr, 118.

Ramos, A., and Fernández, L. 1994. Las pesquerías de merluzas en los caladeros de África noroccidental: datos de base del año 1991. Inf Téc Inst Esp Oceanogr, 153: 132.

Ramos-Martos, A., and Fernández-Peralta, L. 1995. Biology and fisheries of north-west African hakes (*M. merluccius*, *M. senegalensis* and *M. polli*). *In* Hake: Biology, Fisheries and Markets. Fish Fisheries Series, 15. Ed. by Alheit J.and and Pitcher T. J.. Chapman and Hall, London. pp. 89–124.

Ramos, A., Fernández, L., and González, R. 1998. The black hake fishery in the Mauritanian EEZ: analysis about the possible application of 30 cm minimum size. Inf Téc Inst Esp Oceanogr, 173: 40pp.

Rey, J., Fernández-Peralta, L., Esteban, A., García-Cancela, R., Salmerón, F., Puerto, M.A., and Piñeiro, C. 2012. Does otolith macrostructure record environmental or biological events? The case of black hake (*Merluccius polli* and *Merluccius senegalensis*). Fisheries Research, 113:159–172.

Rey, J., Fernández-Peralta, L., Quintanilla, L., Hidalgo, L., Presas, C., Salmerón, F., and Puerto, M. A. 2015. Contrasting energy allocation strategies of two sympatric hake species in an upwelling system. Journal of Fish Biology, 86: 1078–1097.

Rey, J., Fernández-Peralta, L., García, A., Nava, E., Clemente, M. C., Otero, P., Villar, E. I., and Piñeiro, C. 2016. Otolith microstructure analysis reveals differentiated growth histories in sympatric black hakes (*Merluccius polli* and *Merluccius senegalensis*). Fisheries Research, 179: 280–290.

R Core Team. 2014. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna. http://www.R-project.org/.

Sharma, R., Winker, H., Levontin, P., Kell, L., Ovando, D., Palomares, M. L. D., Pinto, C. *et al.*. 2021. Assessing the potential of catch-only models to inform on the state of global fisheries and the UN's SDGs. Sustainability, 13: 6101.

Schnute, J. 1977. Improved estimates from the schaefer production model: theoretical consideration. Journal of the Fisheries Research Board of Canada, 34: 583–603.

Soto, M., Mosqueira, I., Die, D., Gaertner, D., and Pallares, P. 2006. Multi-species assessment of tunas caught in the tropical Atlantic purse seine fishery: sensitivity of production models to biases in reported catche. Collective volume of scientific papers. ICCAT, 59: 546–554.

Su, Yu-Sung, and Yajima, M. 2012. R2jags: A Package for Running Jags from R. R Package Version 0.03-08, Vienna. http://CRAN.R-project.org/package=R2jags (last accessed September 2019).

STECF. 2017. Mediterranean Stock Assessments 2017 part I (STECF-17-15). Publications Office of the European Union, Luxembourg. https://doi.org/10.2760/897559

Thorson, J. T., Ono, K., and Munch, S. B. 2014. A Bayesian approach to identifying and compensating for model misspecification in population models. Ecology Society of America, 95: 329–341.

Thorson, J. T., and Minto, C. 2015. Mixed effects: a unifying framework for statistical modelling in fisheries biology. ICES Journal of Marine Science, 72: 1245–1256.

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

Handling Editor: Sam Subbey