Running head: Managing harvest with biased assessments

Shaping sustainable harvest boundaries for marine populations despite estimation bias

Daisuke Goto[a,d,†], Jennifer A. Devine[a,e], Ibrahim Umar[a], Simon H. Fischer[b], José A. A. De Oliveira[b], Daniel Howell[a], Ernesto Jardim[c,f], Iago Mosqueira[c,g], and Kotaro Ono[a].

[a]Institute of Marine Research/Havforskningsinstituttet, Postboks 1870 Nordnes, 5817 Bergen, Norway

[b]The Centre for Environment, Fisheries and Aquaculture Science (Cefas), Lowestoft Laboratory, Pakefield Road, Lowestoft, Suffolk NR33 0HT, UK

[c]European Commission, DG Joint Research Center, Directorate D – Sustainable Resources, Unit D.02 Water and Marine Resources, Via Enrico Fermi 2749 21027, Ispra, VA, Italy

Present address

[d]Swedish University of Agricultural Sciences, Department of Aquatic Resources, Institute of Freshwater Research, Stångholmsvägen 2, SE-178 93 Drottningholm, Sweden

[e]National Institute of Water & Atmospheric Research, 217 Akersten Street, Nelson 7040, New Zealand

[f]Marine Stewardship Council, Marine House, 1 Snow Hill, London, EC1A 2DH, UK

[g]Wageningen Marine Research, PO Box 68, 1970AB, IJmuiden, The Netherlands

[†]**Corresponding author**: daisuke.goto@slu.se.

Abstract

Systematic errors in population status such as overestimated abundance are pervasive conservation problems. These problems have plagued assessments of commercial exploitation of marine species and can threaten its sustainability. We develop a computer-intensive approach to persistent estimation bias in assessments, which may emerge from unknown sources, by optimizing harvest measures with closed-loop simulation of resource–management feedback systems: management strategy evaluation. Using saithe (*Pollachius virens*), a bottom-water, apex predator in the North Sea, as a real-world case study, we illustrate the approach by first diagnosing robustness of the existing harvest measures and then optimizing the measures through propagation of biases (overestimated stock abundance and underestimated fishing pressure) along with select process and observation uncertainties. Analyses showed that severe biases initially set overly optimistic catch limits and then progressively magnify the amplitude of catch fluctuation, thereby posing unacceptably high overharvest risks. Consistent performance of management strategies can be achieved by developing robust harvest measures that explicitly account for estimation bias through a computational grid search for control parameters that maximize yield while keeping stock abundance above a precautionary level. When the biases become more severe, optimized harvest measures–for saithe, raising threshold abundance that triggers management action and lowering target exploitation rate–would not only safeguard against overharvest risk (<3.5%) but also provide near-term stability in catch limits (<20% year-to-year variation): minimum disruption in fishing communities. The precautionary approach to fine-tuning adaptive risk management through management strategy evaluation offers a powerful tool to better shape sustainable harvest boundaries for exploited marine populations when estimation bias persists. By explicitly accounting for emergent sources of uncertainty, our

proposed approach ensures effective conservation and sustainable exploitation of living marine resources even under profound uncertainty.

Keywords

Decision making; environmental stochasticity; measurement error; management procedure; management strategy evaluation; risk analysis; state-space models; stock assessment; trade-offs.

INTRODUCTION

Managers and policymakers increasingly face trade-offs in sustainably managing extractive use of living marine resources while effectively conserving biodiversity under the precautionary principle (FAO 1996; Hilborn *et al.* 2001; Harwood & Stokes 2003). But imperfect knowledge of social–ecological systems impedes the decision making. Scientific uncertainty of current population status can obscure assessment of decline or extinction threats (Ripa & Lundberg 1996; Ovaskainen & Meerson 2010). Lack of certainty in socioeconomic dynamics that can promote noncompliance and inertia also may reduce the efficacy of management measures applied (Hilborn *et al.* 2001; Beddington, Agnew & Clark 2007; Fulton *et al.* 2011). If we are to achieve internationally agreed conservation targets such as Sustainable Development Goal 14 (UN 2015) and Aichi Biodiversity Target 6 (CBD 2010), we must account for various sources of uncertainty to assess overexploitation risk (Regan *et al.* 2005; Memarzadeh & Boettiger 2018) and recovery potential (Memarzadeh *et al.* 2019) and set conservation priorities.

In commercial capture fisheries, assessments of current population status provide a scientific basis for setting a threshold for safe harvest to prevent the decline of fish stocks by, where necessary, taking precautionary measures. This approach may include the use of biological thresholds such as the population abundance that produces maximum sustainable yield (Beddington, Agnew & Clark 2007). The harvest of wild populations is commonly managed by

applying decision rules based on such predefined thresholds to set a catch limit for the year (Beddington, Agnew & Clark 2007). And accurate population assessments contribute to successful implementation of management measures to sustain long-term commercial exploitation of marine animals (Hilborn *et al.* 2020). But scientific uncertainty in assessments has posed a multitude of challenges (Patterson *et al.* 2001; Sethi 2010). If population abundance is persistently overestimated, for example, resulting overly optimistic catch advice or rebuilding plans will deplete the population, thereby threatening the sustainability of fisheries that depend on it (Walters & Maguire 1996; Memarzadeh *et al.* 2019). Overestimated abundance and underestimated exploitation rate, which often heighten extinction risk, have led to some historical collapses of oceanic predators (Walters & Maguire 1996; Charles 1998).

Systematic errors in perceived population status have plagued assessments of exploited marine species (Punt *et al.* 2020) and likely contributed to overharvest and depletion including stocks that are considered well-monitored (Brooks & Legault 2016). Inconsistency across assessments such as persistent overestimation of abundance has led to the rejection of assessments (Punt *et al.* 2020). Although past research has proposed solutions to estimation bias, applying these solutions remains a challenge because bias could originate from multiple sources (Hurtado-Ferro *et al.* 2015; Brooks & Legault 2016; Szuwalski, Ianelli & Punt 2017). Incomplete knowledge of the causes behind estimation bias a priori may lead to incorrect application of the tools, inadvertently exacerbating the problems by amplifying overharvest and depletion risks (Brooks & Legault 2016; Szuwalski, Ianelli & Punt 2017). Given serious ecological and socioeconomic implications for getting it wrong, we urgently need a procedure that provides practical guidance for explicitly evaluating robustness of management strategies and designing alternative protective measures to

inform decision making to safely harvest marine resources under uncertainty (Kraak *et al.* 2010; Punt *et al.* 2020).

We illustrate how closed-loop simulation of resource–management systems (management strategy evaluation) can help prevent emergent estimation bias from derailing effective management of exploited marine populations. Management strategy evaluation is a flexible decision-support tool used in fisheries management (Butterworth & Punt 1999; Smith, Sainsbury & Stevens 1999) and has increasingly been adopted for conservation planning of imperiled species in marine and terrestrial systems (Milner-Gulland *et al.* 2001; Bunnefeld, Hoshino & Milner-Gulland 2011). This tool is designed to evaluate the performance of candidate policy instruments through forward simulations of feedback (learning from implementation and new observation) between natural resources and management systems by accounting for tradeoffs among management goals of stakeholders (Punt *et al.* 2016). Management strategy evaluation can also assess consequences of suspected sources of bias in assessments (Szuwalski, Ianelli & Punt 2017; Hordyk, Huynh & Carruthers 2019). Here we take the management strategy evaluation approach further: we first diagnose estimation bias (robustness testing, Cooke 1999), and through computational optimization of control parameters of harvest rules (Walters & Hilborn 1978; Chadès *et al.* 2017), our proposed method searches for robust harvest measures by explicitly accounting for emergent bias in perceived stock status along with select process and observation uncertainties. Specifically, we ask: 1) How robust are current management procedures to persistent estimation bias? and 2) How precautionary do management procedures need to be to avert mismanagement–setting overly optimistic catch limits that promote stock depletion and a fishery closure?

METHODS

Management strategy evaluation framework

We simulated annual resource surveys and assessments to explore trade-offs in achieving conservation-oriented (minimizing risk) and harvest-oriented (maximizing yield and minimizing yield variance) goals through management strategy evaluation. We made use of the framework developed and adopted for commercially harvested species in the Northeast Atlantic including four North Sea demersal fish stocks (ICES 2019c) and Atlantic mackerel (*Scomber scombrus*, ICES 2020c). The framework consists of submodels that simulate 1) true population and harvest dynamics at sea and observations through monitoring surveys (operating model, OM), and 2) management processes–assessments based on observations from the surveys and subsequent decision making (management procedure, MP) (Fig. 1a). We used the North Sea population of saithe (*Pollachius virens*) (ICES 2019c), a demersal (bottom-water) predatory fish harvested commercially by more than a dozen European nations, as a real-world case study. We performed all simulations in R (version 3.60, R Development Core Team 2019) using the mse R package (https://github.com/flr/mse) (ICES 2019c), part of the Fisheries Library in R (FLR, Kell *et al.* 2007).

Population dynamics

To simulate future population dynamics of target species, the framework uses an age-structured population model that accounts for environmental stochasticity. For saithe we modeled the population dynamics of juveniles and adults (four-year-olds and older) as

$$\log N_{a,y} = \log N_{a-1,y-1} - F_{a-1,y-1} - M_{a-1,y-1} + \eta_{a,y} \tag{1a}$$

$$\log N_{A,y} = \log(N_{A-1,y-1}e^{-F_{A-1,y-1}-M_{A-1,y-1}} + N_{A,y-1}e^{-F_{A,y-1}-M_{A,y-1}}) + \eta_{A,y} \tag{1b}$$

$$\log F_{a,y} = \log F_{a,y-1} + \xi_{a,y} \tag{1c}$$

where $N_{a,y}$, $F_{a,y}$, and $M_{a,y}$, are a-year-old numbers, fishing mortality rates, and natural mortality (non-fishing such as starvation and diseases) rates in year y, and η and ξ are normally distributed variables, reflecting measurement errors (Nielsen & Berg 2014). Historical surveys indicate that 10-year-olds and older are relatively uncommon, and we simulated them as a dynamic aggregate pool (N_A, F_A, and M_A). We conditioned the OM on the 2018 assessment (ICES statistical areas: Subareas 4 and 6 and Division 3a, Fig. 1c and (ICES 2018), which represents the best available information on the past (1967–2017) population and harvest dynamics (Fig 1c and Appendix S1). We projected forecasts annually for 21 years (2018–2038). The data sources, survey methods, and model structure have been extensively documented in ICES (2016) and ICES (2019a). Briefly, we parameterized the model with 51-year estimates of age-specific masses (g, Appendix S1: Table S3–S5) and maturity rates (proportion of adults, Appendix S1: Table S6), and natural mortality rates assumed at a value of 0.2 year^{-1} for all ages and years. Then, we fitted the population model to time series data of commercial catch (age-aggregated biomass of German, French, and Norwegian trawlers in 2000–2017, tonnes or t, Appendix S1: Table S7) and age-specific (ages three to eight) abundance indices (International bottom trawl surveys in the third quarter, IBTS-Q3, in 1992–2017, Appendix S1: Table S8) (ICES 2018).

To account for process uncertainty (year-to-year variability in survival rate), we generated 1000 realizations of stochastic populations using the variance-covariance matrix of estimable parameters (age-specific numbers and fishing mortality rates) taken from the 2018 assessment (Appendix S2: Fig. S1a, ICES 2019c). We derived a set of mean age-specific masses, maturity rates, and fishing gear selectivity by randomly selecting a year with replacement from the 2008–2017 data; this process was repeated independently for each replicate every simulation year to account for environmental stochasticity.

We simulated density-dependent regulation of recruitment with a segmented regression (ICES 2019c) relating adult biomass to the number of recruits (three-year-olds for saithe). Adult biomass (spawning stock biomass, SSB, t) is the product of age-specific numbers, masses, and maturity rates. We parameterized the spawner–recruit model by fitting it to the 1998–2017 data on SSB and recruit numbers. To account for environmental stochasticity in density-dependency of recruitment, we first used a kernel density function to smooth the resulting distribution of residuals from the fitted regression. Then, we resampled residuals with replacement from the distribution and applied to model outputs to generate recruits every simulation year (Appendix S2: Fig. S2a,b); this process was repeated independently for each replicate. Preliminary analyses showed little evidence of temporal autocorrelation in recruitment (Appendix S2: Fig. S2c).

Monitoring and catch surveys

We simulated future annual monitoring of the population and harvest, which are subject to error and bias, by adding observation error to simulated-true survey indices and age-specific catch computed from the population OM. In forecasting we assumed the model is fixed (averages of life-history parameters such as maturity rates are time-invariant). We simulated deviances to the observed survey index (IBTS-Q3) using the variance-covariance matrix for the survey index to account for observation error correlated between ages (Appendix S2: Fig. S3a). Survey observations (I) are generated as:

$$I_{a,y} = q_a N_{a,y} e^{-t_i Z_{a,y}}\, e^{\varepsilon_{a,y}} \tag{2a}$$

$$\varepsilon_{a,y,i} \sim N(0, \Sigma_i) \tag{2b}$$

where $N_{a,y}$ and $Z_{a,y}$ are a-year-old numbers and total ($F + M$) mortality rates in year y from the population OM; q_a are a-year-old survey catchabilities for the survey i; t is the timing of the annual survey (0.575 for IBTS-Q3). $\varepsilon_{a,y}$ represents multivariate normally distributed errors with

mean zero and standard deviation Σ defined by the variance-covariance matrix between ages within years (ICES 2019b). Observation error is included on age-specific abundance indices as multiplicative lognormal error (Appendix S2: Fig. S3a).

We simulated uncertainty in reported catch by computing a commercial catch index (or exploitable biomass index) generated from the population OM (Appendix S2: Fig. S3b) as:

$$I_y = q\left[\sum_a S_{a,y} w^c_{a,y} N_{a,y} e^{-0.5\, Z_{a,y}}\right] e^{\varepsilon_y} \tag{3a}$$

$$S_{a,y} = F_{a,y} / \sum_a F_{a,y} \tag{3b}$$

$$\varepsilon_y \sim N(0, \sigma^2) \tag{3c}$$

where $N_{a,y}$ and $Z_{a,y}$ are as above; q is the catchability of the commercial catch index; $w^c_{a,y}$ are a-year-old catch masses in year y; 0.5 indicates projection to mid-year; S is the relative F for a-year-olds in year y; and ε_y is a normally distributed error with mean zero and standard deviation σ in year y (Appendix S2: Fig. S1c).

Management procedure

The MP simulates decision making by managers based on perceived current stock status and model-based harvest rules (Fig. 1a). The current status is assessed annually by fitting an estimation model (EM) to the time series (past plus most recent) data passed on from the observation model (survey and catch indices, $I_{a,y}$ and I_y) before provision of catch advice (May for saithe). In this study, we used the State-space Assessment Model (Nielsen & Berg 2014) as an EM and harvest rules set for saithe (ICES 2019c); model settings and forecast assumptions are fully described in ICES (2019c). Under the harvest rules, the following year's catch limit is the product of target exploitation rate (F_{target}) and stock abundance (t) when the estimated SSB in the current assessment year (terminal year) remains above a fixed threshold ($B_{trigger}$) (Fig. 1b). These two control parameters of the decision model are designed to prevent overharvesting by

accounting for uncertainty in population and harvest dynamics (Rindorf *et al.* 2016). When the SSB falls below $B_{trigger}$, exploitation rate is adjusted to F_{target} scaled to the proportion of SSB relative to $B_{trigger}$ (Fig. 1b), thereby allowing the population to rebuild–adaptive harvesting.

Population and management measure performance

We computed conservation-oriented (risk) and harvest-oriented (median catch and interannual catch variability, ICV) metrics from 1000 realizations of short-term (2019–2023) and long-term (2029–2038) projections from the OM to evaluate performance of the harvest rules applied. We chose the number of replicates based on the stability of risk (ICES 2019c). Risk (the probability of quasi-extinction) is defined here as the maximum annual probability of SSB falling below a limit threshold, B_{lim} (Fig. 1b), a spawner abundance below which reproductive capacity of the population is expected to decline (Rindorf *et al.* 2016), consistent with previous analyses (ICES 2019b). We computed the risk based on the proportion of 1000 simulations with annual estimates of SSB < B_{lim}. We estimated B_{lim} using the Eqsim R package (https://github.com/ices-tools-prod/msy); B_{lim} is set to 107 297 t for saithe (Fig. 1b, ICES 2019c). We computed ICV (a percentage change in catch limits) as

$$\mathrm{ICV}_y = \frac{|C_{y+1} - C_y|}{C_y} \times 100 \tag{4}$$

where C_{y+1} and C_y are projected catches in year y+1 and y (respectively).

Estimation bias scenarios

To evaluate how managing with persistent estimation bias degrades performance of harvest rules and potential to achieve management goals, we simulated hypothetical scenarios of bias in perceived spawner abundance and fishing mortality rate in annual assessments. Although bias can emerge in both directions (over- and under-estimation), they have asymmetric implications for conservation and harvest decision making by managers (Hordyk, Huynh & Carruthers 2019).

We analyzed scenarios that can cause severe conservation issues for exploited species: SSB overestimation and mean *F* underestimation simultaneously. We simulated six scenarios by introducing a bias (0%/baseline, 10%, 20%, 30, 40%, and 50% per year) in estimating age-specific numbers in the terminal year of annual assessments before forecasting SSB and mean *F* and projecting catch limits. The magnitudes of realized biases in these parameters varied among simulations because of process uncertainty. Note that for simplicity and generality these bias scenarios are designed to illustrate our proposed approach to generic, systematic estimation error in assessments, rather than specific scenarios of persistent, time-varying bias that may cumulatively emerge between assessments as input data are updated owning to model misspecification and biased input data (known as retrospective pattern, ICES 2020b; Punt *et al.* 2020).

We considered a harvest rule evaluated for North Sea saithe in the latest management strategy evaluation (Fig. 1b, ICES 2019c). For consistency we used the same decision model control parameter values (F_{target} and $B_{trigger}$) that had been estimated through computational optimization by maximizing median catch limits while meeting the precautionary criterion: long-term risk ≤ 0.05 (Fig. 1b, ICES 2019c). We analyzed all scenarios based on the performance metrics (risk, median catch, and ICV) from 1000 realizations of short-term and long-term projections.

Developing robust management measures

To evaluate how precautionary the harvest rule needs to be to minimize disruption in catch advice provisioning when the assessment is biased, we explored alternative values of the two parameters of the decision model ($B_{trigger}$ and F_{target}) and projected catch limits under the same bias scenarios (overestimated SSB and underestimated mean *F*) through management strategy evaluation. Building on the grid search from the latest evaluation (ICES 2019c), in which the

control parameters had been optimized under the baseline scenario, we explored a finite number of select candidate combinations of the parameters (192 per scenario or 1920 unique runs in total) for reopimization to illustrate our proposed approach. We conducted a restricted grid search in parameter space of $B_{trigger}$ (210,000 to 320,000 t with 10,000 t increments) and F_{target} (0.24 to 0.39 with 0.01 increments). We computed median catch limits and risk from 1000 realizations of 21-year simulations. For consistency we used the same precautionary criterion for optimizing the parameter sets: maximizing median catch limits while maintaining long-term risk ≤ 0.05 (ICES 2019b).

RESULTS

Performance of harvest measures with estimation bias

An increasing amount of estimation bias in annual assessments was projected to increase median catch and overharvest risk in the short term. Although median SSBs declined by as much as 30% in the OM (Fig. 2a), with SSB overestimation, median catches rose by 15–44% relative to the baseline (Fig. 3a), increasing mean *F*s in the OM by 19–80%, which were underestimated in the EM by on average 42% (Fig. 2b). As a result, biased assessments elevated risks as much as 17-fold (Fig. 3a). Mean ICV responded nonlinearly to biased estimates, and the distribution was highly skewed (Fig. 3a).

In the long-term, the estimation bias was projected to increase ICV and risk but have negligible effect on median catch. Biased estimates reduced median SSB in the OM by as much as 35% (resulting in a 37% increase in mean *F*) relative to the baseline; this reduction was underestimated in the EM by on average 53% (Fig. 2a,b). With overestimated SSBs and largely unadjusted F_{target}, median catches remained unchanged (~113,000 t, Fig. 3b). Also, biased assessments amplified temporal variations in both SSB and mean *F* in the OM as much as ~71%,

thereby increasing ICVs by up to 72%, which, combined with reduced SSBs, elevated risks 2–13-fold (Fig. 3b).

Harvest rule optimization

The proportion of the select grid search area evaluated through management strategy evaluation that remained precautionary (safe harvest margin) progressively shrank as more bias was introduced (Fig. 4 and Table 1). Within the safe harvest margin, the fishery yielded highest catches at lower (by 0.02–0.10) F_{target} and higher (by 10,000–60,000 t) $B_{trigger}$ (Table 1 and Fig. 4). With reoptimization of these control parameters, the harvest rule was projected to produce higher (by 6.7–25%) short-term catches and maintain similar (<3.0% deviation from the baseline) long-term catches under all bias scenarios (Table 1). And both short- and long-term SSBs declined by 3.1–6.9% and long-term ICVs rose by less than 1.5% (Table 1).

DISCUSSION

A precautionary approach applied through management strategy evaluation offers a powerful decision-support tool to develop robust measures for sustainable harvest of marine populations even when severe estimation bias persists in assessments. For North Sea saithe, increasingly severe biases (abundance overestimation and fishing pressure underestimation) initially set overly optimistic catch limits that deplete the stock. But unacceptably high long-term risks of missing management targets result from progressively amplified fluctuations in annual catch limits. With computational optimization of control parameters, our proposed approach can help develop harvest measures to achieve robust, cost-effective performance: low risks and stable catch limits–less disruption in fishing communities. By explicitly accounting for emergent uncertainty such as persistent estimation bias in stock abundance, this approach can guide

decision makers in balancing common trade-offs in managing commercial exploitation: achieving socioeconomic goals while also conserving living marine resources.

Costs of managing with estimation bias

How robust management measures are to estimation errors in assessment would depend on life history, fishing operation, and current status of a given species or population (Hurtado-Ferro *et al.* 2015; Wiedenmann & Jensen 2018; Hordyk, Huynh & Carruthers 2019). Our North Sea saithe case study is conditioned on the 2018 assessment in which the stock is in good condition (~37% above the biomass that could produce maximum sustainable yield or MSY $B_{trigger}$, ICES 2019c). Analyses show the current harvest rule is robust to a moderate amount of bias (up to ~16%) in assessments and the stock can be sustainably managed at an acceptable level of risk (≤5%). Simulations revealed, however, that managing harvest with more severely biased assessments can progressively amplify the risk of overharvesting (quasi-extinction) but the causes of heightened risk vary over time. The risk initially increases as the population becomes depleted owing primarily to overly optimistic projections of annual catch limits. Past research suggests that this pattern can emerge from misspecification of an estimation model such as unaccounted temporal variability in demographic parameters (Szuwalski, Ianelli & Punt 2017) and underestimated natural mortality rate (Hordyk, Huynh & Carruthers 2019), and biased input data such as underreported catch (Hordyk, Huynh & Carruthers 2019). Over time managing with biased assessments would destabilize the stock, which is displayed as amplified variations in both stock abundance and fishing pressure in our case study. As a result, yields also would become increasingly more variable (by as much as 74% for saithe), elevating the probability of overharvesting. Even when the long-term risk of managing with estimation bias remains

acceptable levels (under <20% bias scenarios in our case study), thus, harvesting destabilized stocks may have more uncertain consequences for population persistence and yield.

Large year-to-year fluctuations in catch limit are disfavored by fishing communities (Anderies 2015) and a management measure to suppress the fluctuations (known as stability or catch constraint) is commonly applied in industrial exploitation (ICES 2019b). But evidence for the efficacy of this policy tool remains limited (but see Kell *et al.* 2005; Kell *et al.* 2006) especially when assessments suggest persistent biases in stock status. Applying the fluctuation-suppressing measure may, to some extent, limit catch variability inflated by managing with biased assessments. But quasi-extinction risks likely remain unacceptably high because this tool may not be sufficiently sensitive to rapid population declines and unlikely prompts large enough reductions in annual catch limits effectively (Kell *et al.* 2005; Kell *et al.* 2006).

The time-varying consequences of estimation bias in assessments also may present a dilemma for managers in decision making, as illustrated for several exploited marine species (Deroba 2014b; Hordyk, Huynh & Carruthers 2019). Managing with biased assessments would produce higher yields (and revenues) in the short term but would amplify catch fluctuations and thus probabilities of depletion and quasi-extinction in the long term. Trade-offs between short-term gains and long-term losses (or vice versa) are common dilemmas in managing natural resources (Mangel *et al.* 1996; Carpenter *et al.* 2015). Past research often focuses on and developing solutions to scientific uncertainty in fisheries management such as assessment error (Brooks & Legault 2016; Wiedenmann & Jensen 2018). But capturing how managers and fishing communities respond to these changes also would contribute to developing effective strategies for sustainable use of living marine resources (Fulton *et al.* 2011). For example, historical records tell us that realized catch limits and landings in the Northeast Atlantic on average varied

less than recommended by scientific advice (Patterson & Résimont 2007), which may attenuate or amplify the effects of biased assessments on the sustainability of harvesting. In situations where science that management advice is based on becomes increasingly unreliable, evaluating both short- and long-term consequences of taking certain management actions would aid managers make decisions effectively. Our findings reemphasize alternative protective measures need to be explicitly assessed before implementation when giving a scientific basis to inform defensible decision making.

Managing risks under rising uncertainty

Our analyses suggest persistent overestimation of abundance and underestimation of fishing pressure can mask the extent of overharvesting and depletion, thereby delaying management responses (asynchronized resource–fishery dynamics, Fryxell *et al.* 2010). Although a certain time lag in management cycle (from monitoring surveys to provisioning of catch advice) is unavoidable, severe estimation bias can promote management inertia. Once population abundance reaches a biological limit threshold (B_{lim}, for example), the population may even become unresponsive to any measure for recovery (Allee effect, Kuparinen, Keith & Hutchings 2014). One proposal to minimize adverse effects of estimation bias is by identifying the sources of and correcting for model misspecification such as accounting for time-varying demographic parameters in an estimation model (Szuwalski, Ianelli & Punt 2017). But without prior knowledge of true demographic processes of the population, the current form of this method may not sufficiently reduce bias or may even exacerbate the problem if incorrectly applied (Szuwalski, Ianelli & Punt 2017). Also, if biases originate from two or more demographic parameters, uncertainties in these misspecified parameters may covary and interact

unpredictably, making the application of the method challenging for many harvested populations.

To circumvent this challenge others suggest annual catch limits be proportionally adjusted using an index that quantifies relative deviation in population metrics (such as stock abundance) between assessments (known as Mohn's ρ) (Deroba 2014b; Brooks & Legault 2016). Although this index can be useful as a diagnostic, past analyses suggest the index may not necessarily reflect the magnitude and direction of estimation bias (Hurtado-Ferro *et al.* 2015; Brooks & Legault 2016; Wiedenmann & Jensen 2018). As such when applied, the outcomes and net benefits can be equivocal in both the short- and long-terms (Deroba 2014b; Brooks & Legault 2016).

Shifting the focus from assessment to decision making in management strategy evaluation (Fig. 1a), our analysis shows the undesirable outcomes of managing with biased assessments can be avoided by developing more precautionary measures to set annual catch limits through dynamic optimization of the control parameters of harvest rules. For our saithe case when estimation bias becomes too severe, lowering target exploitation rate and raising threshold abundance that trigger management action–early intervention–would maintain not only low probabilities of quasi-extinction (<3.5%) (and thus a fishery closure) but also near-term catch stability (<20% year-to-year variation) without foregoing yields, thereby minimizing disruption in fishing communities. Although this approach needs to be tested with more case studies, our work demonstrates the optimization approach can guide managers in making decisions cost-effectively to safeguard against ecologically and socioeconomically undesirable outcomes of managing risks under systematic uncertainty.

Like all model-based methods our proposed approach also has limitations. Our main aim of this work was to develop an alternative approach to guide resource managers in decision making to support the sustainable use of living marine resources despite systematic estimation error. For this reason, we did not explore underlying causes of the estimation error propagating through a resource–management system. But analyses show that even with optimization our ability to safely harvest marine resources would become progressively limited (less margin of error in setting the precautionary harvest rules or "safe operating space", Carpenter *et al.* 2015) as the magnitude of estimation bias increases. Thus, we encourage continued efforts to develop methods to identify root causes of estimation bias and to minimize their adverse effects on scientific advice (Hurtado-Ferro *et al.* 2015; Szuwalski, Ianelli & Punt 2017; Hordyk, Huynh & Carruthers 2019).

Another caveat of our approach is computational intensity (requiring extensive parallel computing on a high-performance computer cluster), which may pose challenges in its application especially for more complex management objectives (more control parameters) (Walters & Hilborn 1978; Chadès *et al.* 2017). But several methods have been recently adopted to improve the efficiency of computational optimization including genetic algorithms (Fischer *et al.* 2021), partially observable Markov decision process (Memarzadeh & Boettiger 2018), bootstrapping (ICES 2020a), and Bayesian statistics (ICES 2020a). Future research would benefit from applying these techniques to expand this feedback-based approach to tackling estimation bias in assessment.

More broadly, the management strategy evaluation approach offers an alternative, robust solution to estimation bias in assessment of managed populations. This approach can not only act as a diagnostic tool to evaluate the robustness of management measures by explicitly accounting

for long-term consequences but also present an adaptive, transparent way to improve protective measures when the perception deviates too far from the reality (when science fails to deliver). Given ubiquity of estimation bias and challenges in identifying the sources (Hurtado-Ferro *et al.* 2015; Brooks & Legault 2016; Szuwalski, Ianelli & Punt 2017), we suggest the bias be routinely evaluated through management strategy evaluation as an additional source of uncertainty, and the control parameters of harvest rules be (re)optimized when the bias becomes too severe.

Demand for wild-capture fisheries, which provide food, nutrition, and job security, will continue to rise with growing human populations in the coming decades (Costello et al. 2020). Changing ocean conditions are also projected to increase environmental stochasticity, amplifying marine population and harvest fluctuations (Brooks & Legault 2016). Higher environmental stochasticity may promote autocorrelation in population fluctuation (Ripa & Lundberg 1996; Gamelon, Sandercock & Sæther 2019) and amplify the magnitude of scientific uncertainty, thereby further shrinking safe harvest margins. These anticipated issues underscore greater needs for taking precautionary measures in shaping resilient management policies (adopting "resilience-thinking", Fischer *et al.* 2009) to safeguard shared living resources in the face of rising uncertainty.

Acknowledgements

We thank all participants of the ICES Workshop of North Sea Management Strategies Evaluation (WKNSMSE) for feedback on the saithe management strategy evaluation work. We also thank Chris Legault and anonymous reviewers for comments on earlier versions of the manuscript. Some figures use images from the IAN Symbols, courtesy of the Integration and Application Network, University of Maryland Center for Environmental Science

(ian.umces.edu/symbols/). This project was partially funded by Institute of Marine Research's REDUS (Reduced Uncertainty in Stock Assessments) project.

Literature Cited

Anderies, J.M. (2015) Managing variance: key policy challenges for the Anthropocene. *P Natl Acad Sci,* **112,** 14402-14403.

Beddington, J.R., Agnew, D.J. & Clark, C.W. (2007) Current problems in the management of marine fisheries. *science,* **316,** 1713-1716.

Brooks, E.N. & Legault, C.M. (2016) Retrospective forecasting—evaluating performance of stock projections for New England groundfish stocks. *Can J Fish Aquat Sci,* **73,** 935-950.

Bunnefeld, N., Hoshino, E. & Milner-Gulland, E.J. (2011) Management strategy evaluation: a powerful tool for conservation? *Trends Ecol Evol,* **26,** 441-447.

Butterworth, D. & Punt, A. (1999) Experiences in the evaluation and implementation of management procedures. *ICES J Mar Sci,* **56,** 985-998.

Carpenter, S.R., Brock, W.A., Folke, C., Van Nes, E.H. & Scheffer, M. (2015) Allowing variance may enlarge the safe operating space for exploited ecosystems. *P Natl Acad Sci,* **112,** 14384-14389.

CBD (2010) The strategic plan for biodiversity 2011–2020 and the Aichi Biodiversity Targets. COP 10 Decision X/2. CBD, Montreal, Canada.

Chadès, I., Nicol, S., Rout, T.M., Péron, M., Dujardin, Y., Pichancourt, J.-B., Hastings, A. & Hauser, C.E. (2017) Optimization methods to solve adaptive management problems. *Theor Ecol,* **10,** 1-20.

Charles, A.T. (1998) Living with uncertainty in fisheries: analytical methods, management priorities and the Canadian groundfishery experience. *Fish Res,* **37,** 37-50.

Cooke, J. (1999) Improvement of fishery-management advice through simulation testing of harvest algorithms. *ICES Journal of Marine Science,* **56,** 797-810.

Deroba, J.J. (2014a) Evaluating the consequences of adjusting fish stock assessment estimates of biomass for retrospective patterns using Mohn's Rho. *North American Journal of Fisheries Management,* **34,** 380-390.

Deroba, J.J. (2014b) Evaluating the Consequences of Adjusting Fish Stock Assessment Estimates of Biomass for Retrospective Patterns using Mohn's Rho. *N Am J Fish Manage,* **34,** 380-390.

FAO (1996) Precautionary approach to capture fisheries and species introductions. FAO Technical Guidelines for Responsible Fisheries 2, FAO.

Fischer, J., Peterson, G.D., Gardner, T.A., Gordon, L.J., Fazey, I., Elmqvist, T., Felton, A., Folke, C. & Dovers, S. (2009) Integrating resilience thinking and optimisation for conservation. *Trends Ecol Evol,* **24,** 549-554.

Fischer, S.H., De Oliveira, J.A.A., Mumford, J.D. & Kell, L.T. (2021) Using a genetic algorithm to optimize a data-limited catch rule. *ICES J Mar Sci*.

Fryxell, J.M., Packer, C., McCann, K., Solberg, E.J. & Sæther, B.-E. (2010) Resource management cycles and the sustainability of harvested wildlife populations. *science,* **328,** 903-906.

Fulton, E.A., Smith, A.D., Smith, D.C. & van Putten, I.E. (2011) Human behaviour: the key source of uncertainty in fisheries management. *Fish Fish,* **12,** 2-17.

Gamelon, M., Sandercock, B.K. & Sæther, B.E. (2019) Does harvesting amplify environmentally induced population fluctuations over time in marine and terrestrial species? *J Appl Ecol,* **56,** 2186-2194.

Harwood, J. & Stokes, K. (2003) Coping with uncertainty in ecological advice: lessons from fisheries. *Trends Ecol Evol*, **18,** 617-622.

Hilborn, R., Amoroso, R.O., Anderson, C.M., Baum, J.K., Branch, T.A., Costello, C., de Moor, C.L., Faraj, A., Hively, D. & Jensen, O.P. (2020) Effective fisheries management instrumental in improving fish stock status. *P Natl Acad Sci*.

Hilborn, R., Maguire, J.-J., Parma, A.M. & Rosenberg, A.A. (2001) The precautionary approach and risk management: can they increase the probability of successes in fishery management? *Can J Fish Aquat Sci*, **58,** 99-107.

Hordyk, A.R., Huynh, Q.C. & Carruthers, T.R. (2019) Misspecification in stock assessments: Common uncertainties and asymmetric risks. *Fish Fish*, **20,** 888-902.

Hurtado-Ferro, F., Szuwalski, C.S., Valero, J.L., Anderson, S.C., Cunningham, C.J., Johnson, K.F., Licandeo, R., McGilliard, C.R., Monnahan, C.C. & Muradian, M.L. (2015) Looking in the rear-view mirror: bias and retrospective patterns in integrated, age-structured stock assessment models. *ICES J Mar Sci*, **72,** 99-110.

ICES (2016) Report of the Benchmark Workshop on North Sea Stocks (WKNSEA), 14–18 March 2016, Copenhagen, Denmark. ICES CM 2016/ACOM:37 704 pp.

ICES (2018) Report of the Working Group on the Assessment of Demersal Stocks in the North Sea and Skagerrak (WGNSSK), 24 April - 3 May 2018, Oostende, Belgium. ICES CM 2018/ACOM:22. 1264 pp.

ICES (2019a) Report of the Interbenchmark Protocol on North Sea Saithe. (IBPNSsaithe). ICES Scientific Reports. VOL 1:ISS 1. 65 pp.

ICES (2019b) Workshop on guidelines for management strategy evaluations (WKGMSE2). ICES Scientific Reports. 1:33. 162 pp.

ICES (2019c) Workshop on north sea stocks management strategy evaluation (WKNSMSE). ICES Scientific Reports. 1:12. 378 pp.

ICES (2020a) The third Workshop on Guidelines for Management Strategy Evaluations (WKGMSE3). ICES Scientific Reports. 2:116. 112 pp. http://doi.org/10.17895/ices.pub.7627.

ICES (2020b) Workshop on Catch Forecast from Biased Assessments (WKFORBIAS; outputs from 2019 meeting). ICES Scientific Reports. 2:28. 38 pp. http://doi.org/10.17895/ices.pub.5997.

ICES (2020c) Workshop on Management Strategy Evaluation of Mackerel (WKMSEMAC). ICES Scientific Reports. 2:74. 175 pp. http://doi.org/10.17895/ices.pub.7445.

Kell, L., Pastoors, M., Scott, R., Smith, M., Van Beek, F., O'Brien, C. & Pilling, G. (2005) Evaluation of multiple management objectives for Northeast Atlantic flatfish stocks: sustainability vs. stability of yield. *ICES J Mar Sci*, **62,** 1104-1117.

Kell, L., Pilling, G., Kirkwood, G., Pastoors, M., Mesnil, B., Korsbrekke, K., Abaunza, P., Aps, R., Biseau, A. & Kunzlik, P. (2006) An evaluation of multi-annual management strategies for ICES roundfish stocks. *ICES J Mar Sci*, **63,** 12-24.

Kell, L.T., Mosqueira, I., Grosjean, P., Fromentin, J.-M., Garcia, D., Hillary, R., Jardim, E., Mardle, S., Pastoors, M. & Poos, J. (2007) FLR: an open-source framework for the evaluation and development of management strategies. *ICES J Mar Sci*, **64,** 640-646.

Kraak, S.B., Kelly, C.J., Codling, E.A. & Rogan, E. (2010) On scientists' discomfort in fisheries advisory science: the example of simulation-based fisheries management-strategy evaluations. *Fish Fish*, **11,** 119-132.

Kuparinen, A., Keith, D.M. & Hutchings, J.A. (2014) Allee effect and the uncertainty of population recovery. *Conserv Biol*, **28,** 790-798.

Mangel, M., Talbot, L.M., Meffe, G.K., Agardy, M.T., Alverson, D.L., Barlow, J., Botkin, D.B., Budowski, G., Clark, T. & Cooke, J. (1996) Principles for the conservation of wild living resources. *Ecol Appl*, **6,** 338-362.

Memarzadeh, M. & Boettiger, C. (2018) Adaptive management of ecological systems under partial observability. *Biol Conserv*, **224,** 9-15.

Memarzadeh, M., Britten, G.L., Worm, B. & Boettiger, C. (2019) Rebuilding global fisheries under uncertainty. *P Natl Acad Sci*, **116,** 15985-15990.

Milner-Gulland, E., Shea, K., Possingham, H., Coulson, T. & Wilcox, C. (2001) Competing harvesting strategies in a simulated population under uncertainty. *Anim Conserv*, **4,** 157-167.

Nielsen, A. & Berg, C.W. (2014) Estimation of time-varying selectivity in stock assessments using state-space models. *Fish Res*, **158,** 96-101.

Ovaskainen, O. & Meerson, B. (2010) Stochastic models of population extinction. *Trends Ecol Evol*, **25,** 643-652.

Patterson, K., Cook, R., Darby, C., Gavaris, S., Kell, L., Lewy, P., Mesnil, B., Punt, A., Restrepo, V. & Skagen, D.W. (2001) Estimating uncertainty in fish stock assessment and forecasting. *Fish Fish*, **2,** 125-157.

Patterson, K. & Résimont, M. (2007) Change and stability in landings: the responses of fisheries to scientific advice and TACs. *ICES J Mar Sci*, **64,** 714-717.

Punt, A.E., Butterworth, D.S., de Moor, C.L., De Oliveira, J.A.A. & Haddon, M. (2016) Management strategy evaluation: best practices. *Fish Fish*, **17,** 303-334.

Punt, A.E., Tuck, G.N., Day, J., Canales, C.M., Cope, J.M., de Moor, C.L., De Oliveira, J.A., Dickey-Collas, M., Elvarsson, B.Þ. & Haltuch, M.A. (2020) When are model-based stock assessments rejected for use in management and what happens then? *Fish Res,* **224,** 105465.

Regan, H.M., Ben-Haim, Y., Langford, B., Wilson, W.G., Lundberg, P., Andelman, S.J. & Burgman, M.A. (2005) Robust decision-making under severe uncertainty for conservation management. *Ecol Appl,* **15,** 1471-1477.

Rindorf, A., Cardinale, M., Shephard, S., De Oliveira, J.A., Hjorleifsson, E., Kempf, A., Luzenczyk, A., Millar, C., Miller, D.C. & Needle10, C.L. (2016) Fishing for MSY: can "pretty good yield" ranges be used without impairing recruitment. *ICES J Mar Sci*.

Ripa, J. & Lundberg, P. (1996) Noise colour and the risk of population extinctions. *P Roy Soc Lond B Bio,* **263,** 1751-1753.

Sethi, S.A. (2010) Risk management for fisheries. *Fish Fish,* **11,** 341-365.

Smith, A., Sainsbury, K. & Stevens, R. (1999) Implementing effective fisheries-management systems–management strategy evaluation and the Australian partnership approach. *ICES J Mar Sci,* **56,** 967-979.

Szuwalski, C.S., Ianelli, J.N. & Punt, A.E. (2017) Reducing retrospective patterns in stock assessment and impacts on management performance. *ICES J Mar Sci,* **75,** 596-609.

UN (2015) Transforming our world: the 2030 Agenda for Sustainable Development A/RES/70/1. *Division for Sustainable Development Goals: New York, NY, USA.*

Walters, C. & Maguire, J.-J. (1996) Lessons for stock assessment from the northern cod collapse. *Rev Fish Biol Fisher,* **6,** 125-137.

Walters, C.J. & Hilborn, R. (1978) Ecological optimization and adaptive management. *Annu Rev Ecol Syst*, **9,** 157-188.

Wiedenmann, J. & Jensen, O.P. (2018) Uncertainty in stock assessment estimates for New England groundfish and its impact on achieving target harvest rates. *Can J Fish Aquat Sci*, **75,** 342-356.

Tables

Table 1. Optimized control parameters (F_{target} and $B_{trigger}$)[†] of the harvest rule set for North Sea saithe and performance metrics[‡] from management strategy evaluation under scenarios of varying levels of estimation bias in assessments.

			short-term (2019–2023)				long-term (2029–2038)				
scenario[§]	F_{target}	$B_{trigger}$	catch	ICV	SSB	risk[¶]	catch	ICV	SSB	risk[¶]	SHM[¶]
base	0.35	250000	92464	20	251973	2.0	116700	17.7	292067	1.5	-
10%	0.33	250000	101786	13	238194	3.2	116288	17.8	279135	2.5	84.4
20%	0.31	270000	103545	13	235356	3.3	116154	18.7	274958	3.0	65.6
30%	0.27	310000	93047	20	252123	2.2	115984	18.0	293711	2.2	53.1
40%	0.26	310000	101131	14	240643	2.9	115863	18.4	282929	2.5	37.5
50%	0.25	310000	104943	12	234525	3.3	115730	19.1	274228	2.8	29.2

[†]The model parameters were optimized at the highest median catch while meeting the precautionary criterion: long-term risk ≤ 5% (ICES 2019c).

[‡]The performance was evaluated with short-term and long-term median catch (t), interannual catch variability (%, ICV), median spawning stock biomass (SSB, t), and risk (%).

[§]Scenarios simulate SSB overestimation and mean fishing mortality rate (*F*) underestimation.

[¶]Risk is the maximum probability of SSB falling below B_{lim} (107,297 t) over a given period. Safe harvest margin (SHM) indicates the proportion (%) of the grid-search area with the harvest rules that remain precautionary (Fig. 4).

Figure legends

Figure 1. Management strategy evaluation framework and historical population and harvest dynamics of North Sea saithe. (a) schematic of the management strategy evaluation framework (Fisheries Library in R/Assessment for All or FLR/a4a, redrawn from https://github.com/ejardim) adopted for evaluation of saithe management strategies. (b) harvest control rule evaluated in this study (c) reconstructed saithe population and harvest dynamics taken from the 2018 assessment. In b, blue dashed (horizontal and vertical) lines show the harvest rule parameters (reference points) set for saithe: B_{trigger} = 250,000 t and F_{target} = 0.35 (ICES 2019c).

Figure 2. Short-term (2018–2023) and long-term (years 2029–2038) stock abundance (SSB) and fishing pressure of North Sea saithe from the population operating and estimation models (OM and EM, respectively) under scenarios of varying levels of estimation bias. Violin plots indicate frequency distributions of performance metrics. Horizontal lines (from bottom to top) within the box plots indicates the 25th, 50th, and 75th percentiles; whiskers extend to the largest and smallest values within 1.5x the inter-quartile range (IQR) from the box edges; and black circles indicate the outliers. Fishing mortality rates are computed by averaging across age-specific fishing mortality rates of 4–7-year-olds.

Figure 3. Short-term (2018–2023) and long-term (years 2029–2038) performance of the harvest rule for North Sea saithe under six scenarios of varying levels of estimation bias (overestimation of stock abundance and underestimation of fishing mortality rate). The performance was evaluated with median catch (t), interannual catch variability (ICV, %), median spawner abundance (SSB, t), and risk. Risk is the maximum probability of SSB falling below B_{lim} (107,297 t). Violin plots indicate frequency distributions of performance metrics. Horizontal

lines (from bottom to top) within the box plots indicates the 25th, 50th, and 75th percentiles; whiskers extend to the largest and smallest values within 1.5x the inter-quartile range (IQR) from the box edges; and black circles indicate the outliers. Red horizontal lines indicate median values from the baseline scenario (SSB, catch, and ICV) or the precautionary threshold (risk = 0.05).

Figure 4. Grid search for combination of the decision model (harvest rule) control parameters (F_{target} and $B_{trigger}$) for North Sea saithe under five scenarios of varying levels of estimation bias (overestimation of stock abundance and underestimation of fishing mortality rate). Heat maps indicate median catch for only combinations that meet the precautionary criterion (risk ≤ 5%) in the long term (years 2029–2038). Black rectangles indicate combinations of the model parameters with the highest median catch. Blue circles indicate the parameter sets optimized without estimation bias ($B_{trigger}$ = 250,000 t and F_{target} = 0.35, ICES 2019c).

Figure 1

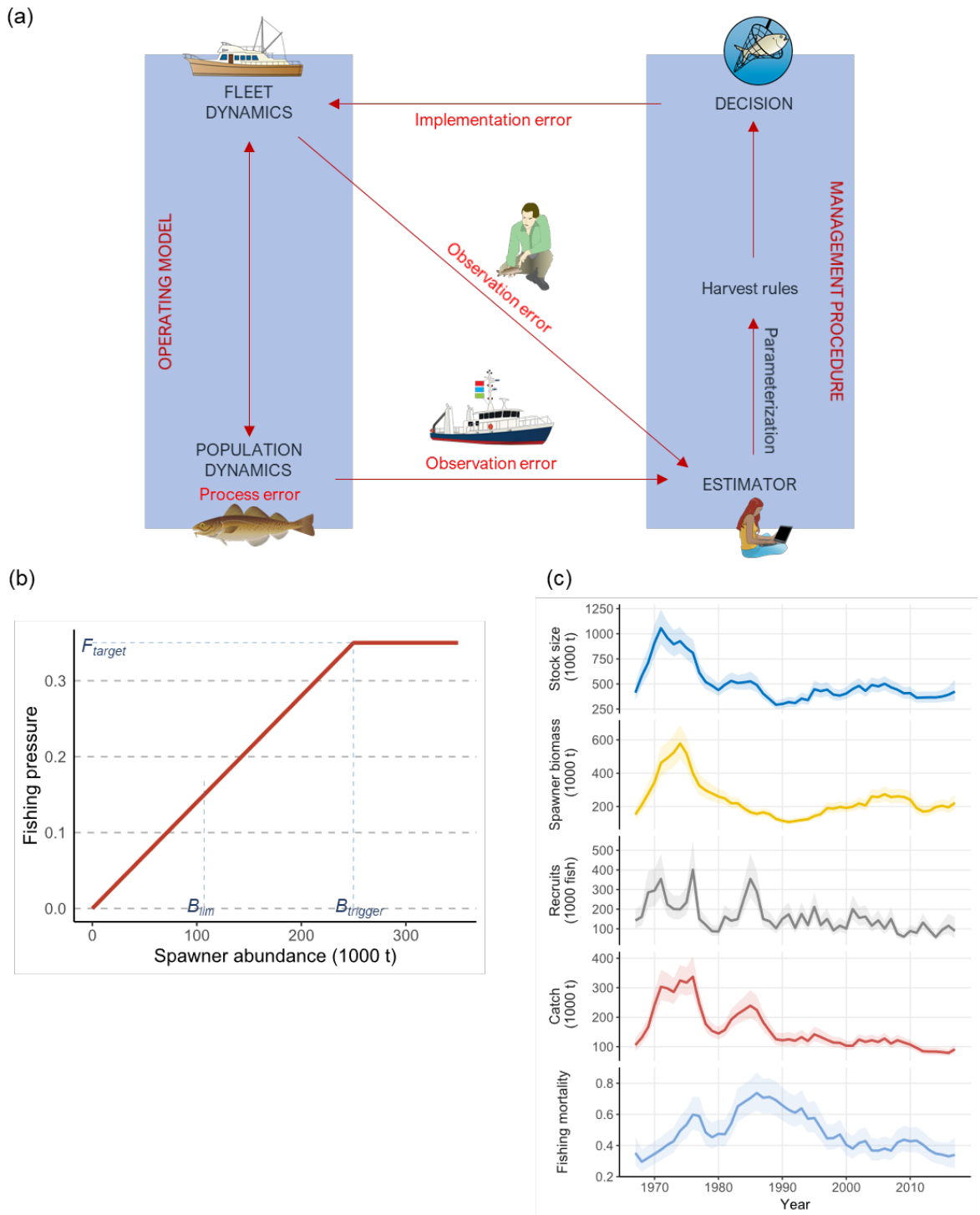

Figure 1. Management strategy evaluation framework and historical population and harvest dynamics of North Sea saithe. (a) schematic of the management strategy evaluation framework (Fisheries Library in R/Assessment for All or FLR/a4a, redrawn from https://github.com/ejardim) adopted for evaluation of saithe management strategies. (b) harvest control rule evaluated in this study (c) reconstructed saithe population and harvest dynamics taken from the 2018 assessment. In b, blue dashed (horizontal and vertical) lines show the harvest rule parameters (reference points) set for saithe: $B_{trigger}$ = 250,000 t and F_{target} = 0.35 (ICES 2019c).

Figure 2.

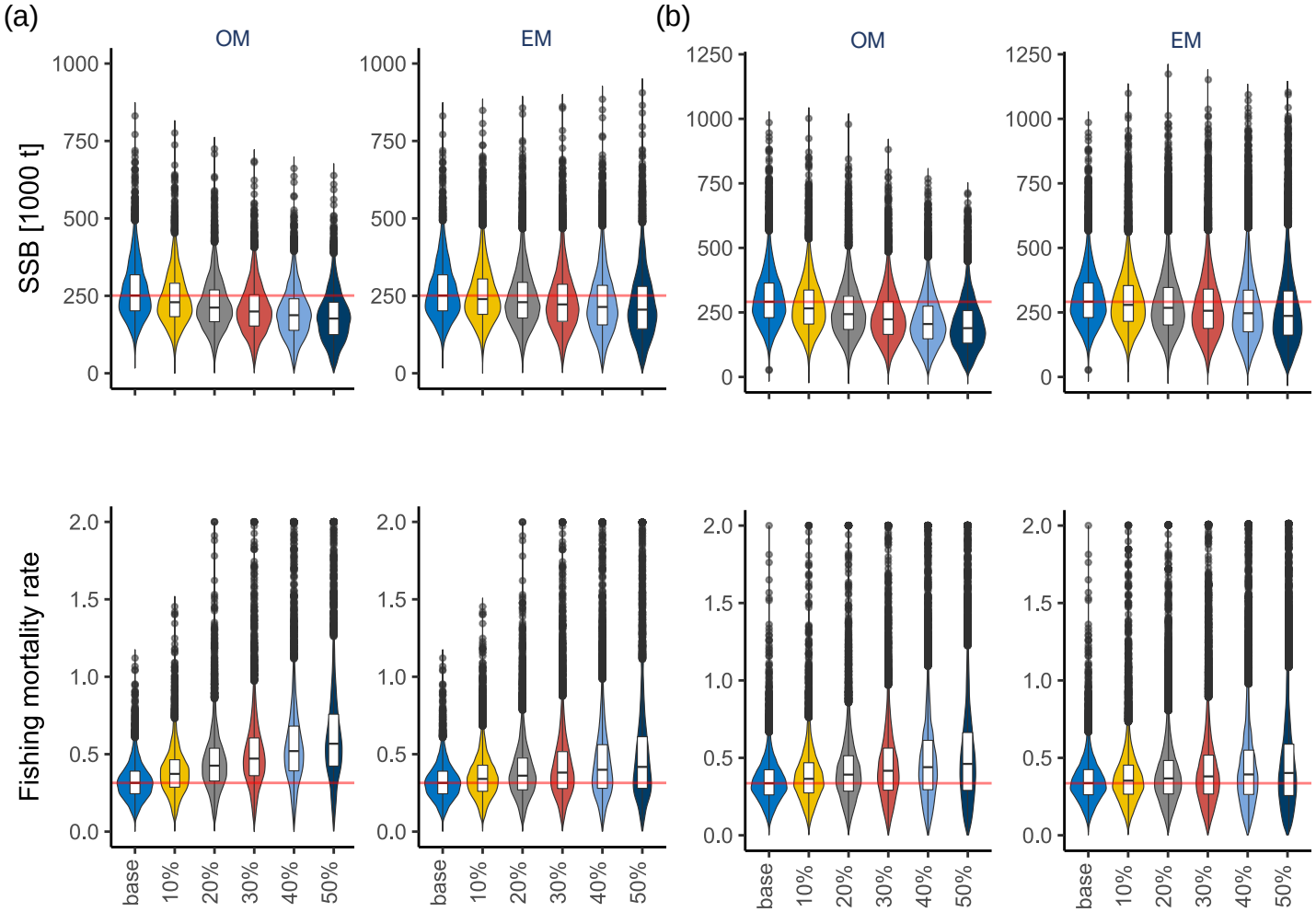

Figure 2. Short-term (years 2018–2023) and long-term (years 2029–2038) stock abundance (SSB) and fishing pressure of North Sea saithe from the population operating and estimation models (OM and EM, respectively) under scenarios of varying levels of estimation bias. Violin plots indicate frequency distributions of performance metrics. Horizontal lines (from bottom to top) within the box plots indicates the 25th, 50th, and 75th percentiles; whiskers extend to the largest and smallest values within 1.5x the inter-quartile range (IQR) from the box edges; and black circles indicate the outliers. Fishing mortality rates are computed by averaging across age-specific fishing mortality rates of 4–7-year-olds.

Figure 3.

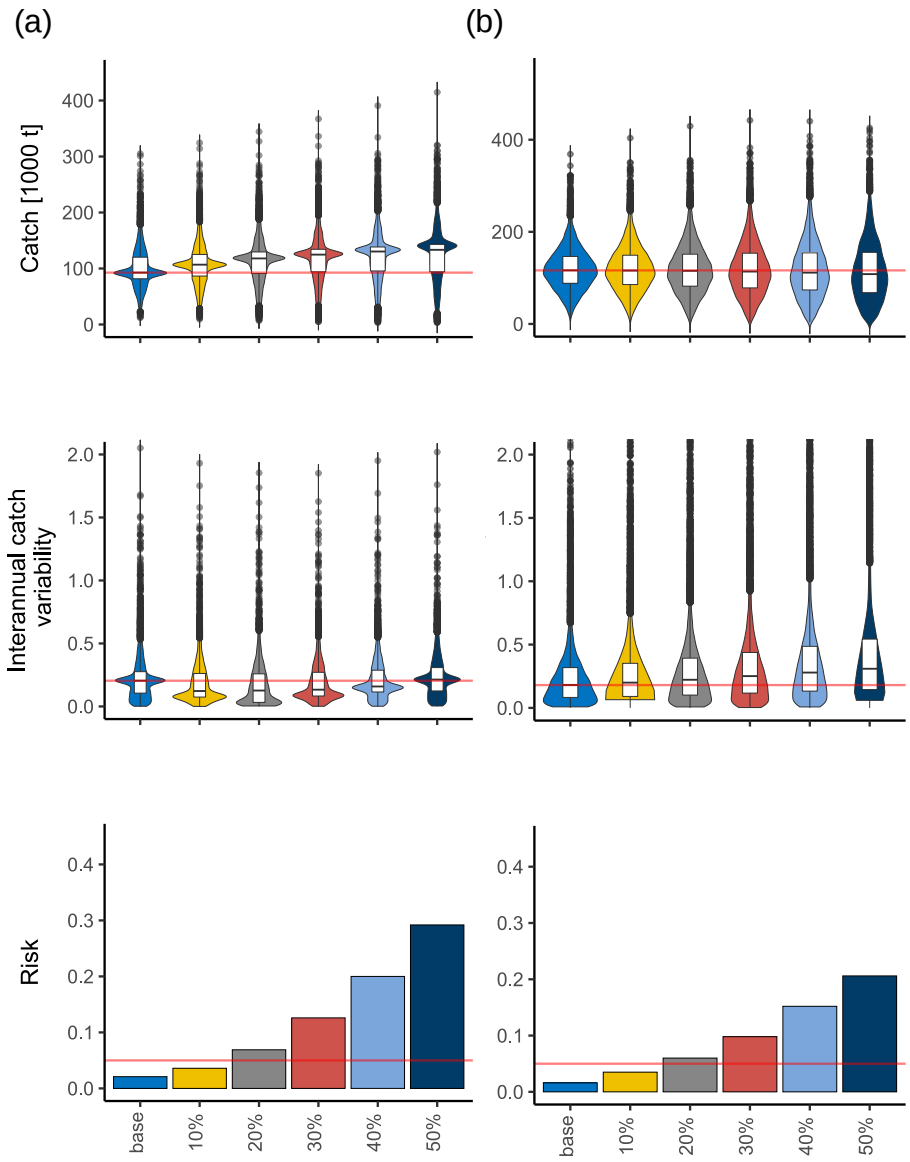

Figure 3. Short-term (years 2018–2023) and long-term (years 2029–2038) performance of the harvest rule for North Sea saithe under six scenarios of varying levels of estimation bias (overestimation of stock abundance and underestimation of fishing mortality rate). The performance was evaluated with median catch (t), interannual catch variability (ICV, %), median spawner abundance (SSB, t), and risk. Risk is the maximum probability of SSB falling below B_{lim} (107,297 t). Violin plots indicate frequency distributions of performance metrics. Horizontal lines (from bottom to top) within the box plots indicates the 25th, 50th, and 75th percentiles; whiskers extend to the largest and smallest values within 1.5x the inter-quartile range (IQR) from the box edges; and black circles indicate the outliers. Red horizontal lines indicate median values from the baseline scenario (SSB, catch, and ICV) or the precautionary threshold (risk = 0.05).

Figure 4.

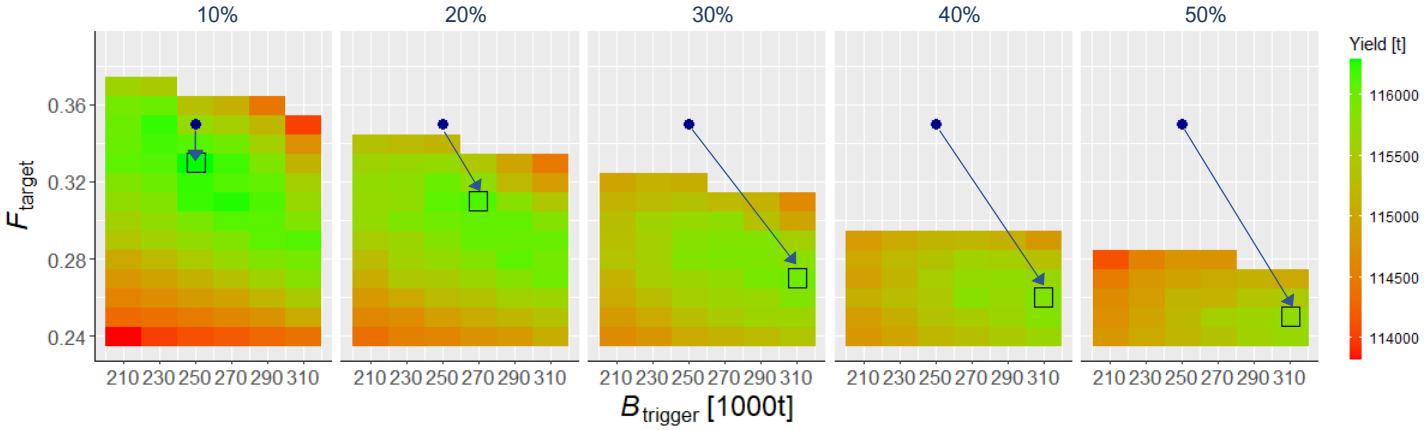

Figure 4. Grid search for combination of the decision model (harvest rule) control parameters (F_{target} and $B_{trigger}$) for North Sea saithe under five scenarios of varying levels of estimation bias (overestimation of stock abundance and underestimation of fishing mortality rate). Heat maps indicate median catch for only combinations that meet the precautionary criterion (risk ≤ 5%) in the long term (years 2029–2038). Black rectangles indicate combinations of the model parameters with the highest median catch. Blue circles indicate the parameter sets optimized without estimation bias ($B_{trigger}$ = 250,000 t and F_{target} = 0.35, ICES 2019c).

Supporting Information. Goto, D., J.A. Devine, I. Umar, S.H. Fischer, J.A.A. De Oliveira, D. Howell, E. Jardim, I. Mosqueira, and K. Ono. Shaping sustainable harvest boundaries for marine populations despite estimation bias.

Appendix S1. Input data and parameter estimates for the North Sea saithe popualtion operating model used in management strategy evaluation

Table S1. Age-specific landings numbers (thousands) of North Sea saithe in 1967–2017.

year	age 3	age 4	age 5	age 6	age 7	age 8	age 9	age 10 and older
1967	17330	16220	15531	2303	1594	292	198	183
1968	23223	21231	13184	6023	429	242	123	145
1969	30235	17681	11057	7609	5738	791	626	150
1970	37249	76661	15000	12128	3894	1792	318	267
1971	69808	57792	32737	4736	4248	2843	1874	774
1972	48075	66095	25317	21207	3672	2944	1641	1607
1973	54332	37698	26849	16061	8428	2000	1357	2381
1974	66938	33740	14123	20688	14666	5199	1477	1955
1975	56987	25864	10319	7566	13657	9357	3501	2687
1976	207823	53060	11696	6253	3976	5362	3586	3490
1977	27461	54967	14755	5490	3777	3447	3812	4701
1978	35059	27269	18062	3312	1138	1033	768	3484
1979	16332	14216	11182	8699	2805	733	540	2089
1980	17494	12341	9015	6718	5658	1150	509	2302
1981	26178	8339	6739	3675	3335	3396	657	2536
1982	31895	40587	9174	5978	2145	1454	982	1254
1983	28242	20604	26013	5678	4893	1494	1036	1327
1984	80933	32172	12957	13011	1657	1252	335	646
1985	134024	55605	13281	4765	3005	682	399	742
1986	55435	91223	15186	5381	2603	1456	445	900
1987	31220	97470	13990	3158	1811	1240	910	700
1988	32578	26408	35323	3828	1908	1104	776	680
1989	22128	30752	13187	10951	1557	739	419	488
1990	40808	19583	11322	4714	2776	745	281	364
1991	46117	29871	7467	3583	1716	953	367	458
1992	18404	33614	12753	3193	1524	696	518	422
1993	37823	20828	11845	3125	1568	1511	814	1026
1994	19958	40193	13034	4297	947	346	427	794
1995	26664	26034	14797	3774	3494	674	552	800
1996	11066	38861	11786	7731	3163	808	210	491
1997	15036	19299	30177	3676	2640	1012	291	288
1998	10363	31017	16367	16077	2231	1206	567	277
1999	9429	13872	26684	8389	10070	2346	891	657
2000	7064	17295	8940	12339	3159	3226	641	441
2001	16052	17646	22421	3349	3586	1772	1614	245
2002	9131	31779	12286	13307	2245	2220	1199	1479
2003	13009	24646	20397	6836	6855	1535	2000	2042
2004	8037	20071	15649	15220	2037	2164	1300	1066
2005	9191	15473	19060	20042	7956	1628	1188	1151
2006	12200	26690	9986	11286	8395	3824	1008	1281
2007	15181	10163	19157	7078	5564	3610	1218	930
2008	6924	23230	10930	14196	4977	3276	3551	3118
2009	6607	14349	13827	5817	8419	2978	1505	2934
2010	7880	8859	9174	6394	2670	4762	1679	2669
2011	10150	22799	8852	3630	2860	1183	1563	3869
2012	7029	11712	15572	4016	1971	1267	537	2610
2013	4999	25516	4974	7645	1886	1241	616	1658
2014	3099	12117	13380	3737	4047	1036	429	1388
2015	6206	7392	13555	8021	1844	1621	715	975
2016	3508	10374	8756	5156	2732	1423	852	1317
2017	3033	15139	8795	6179	5362	1876	820	1111

Table S2. Age-specific discards numbers (thousands) of North Sea saithe in 1967–2017.

year	age 3	age 4	age 5	age 6	age 7	age 8	age 9	age 10 and older
1967	9617	3175	1141	55	16	7	5	2
1968	12888	4156	969	143	4	6	3	2
1969	16779	3461	813	181	57	19	16	2
1970	20671	15007	1102	288	38	42	8	3
1971	38741	11313	2406	112	42	67	48	9
1972	26680	12938	1861	504	36	69	42	18
1973	30152	7380	1973	381	83	47	35	26
1974	37148	6605	1038	491	144	122	38	22
1975	31626	5063	758	180	135	220	89	30
1976	115333	10387	860	148	39	126	92	38
1977	15240	10760	1084	130	37	81	97	52
1978	19456	5338	1327	79	11	24	20	38
1979	9063	2783	822	207	28	17	14	23
1980	9709	2416	662	160	56	27	13	25
1981	14527	1632	495	87	33	80	17	28
1982	17700	7945	674	142	21	34	25	14
1983	15673	4033	1912	135	48	35	26	15
1984	44915	6298	952	309	16	29	9	7
1985	74378	10885	976	113	30	16	10	8
1986	30764	17857	1116	128	26	34	11	10
1987	17326	19080	1028	75	18	29	23	8
1988	18079	5169	2596	91	19	26	20	7
1989	12280	6020	969	260	15	17	11	5
1990	22647	3833	832	112	27	18	7	4
1991	25593	5847	549	85	17	22	9	5
1992	10213	6580	937	76	15	16	13	5
1993	20990	4077	871	74	15	36	21	11
1994	11076	7868	958	102	9	8	11	9
1995	14797	5096	1087	90	34	16	14	9
1996	6141	7607	866	184	31	19	5	5
1997	8344	3778	2218	87	26	24	7	3
1998	5751	6072	1203	382	22	28	14	3
1999	5233	2716	1961	199	99	55	23	7
2000	3920	3386	657	293	31	76	16	5
2001	8908	3454	1648	80	35	42	41	3
2002	8439	5710	2451	425	64	324	121	96
2003	15288	7106	234	0	0	0	0	0
2004	5605	4407	0	0	0	0	0	0
2005	3498	0	0	0	0	0	0	0
2006	5114	5282	394	0	0	0	0	0
2007	9433	3152	1762	97	0	0	0	0
2008	696	7682	1610	745	111	9	0	0
2009	831	1158	395	30	93	16	14	11
2010	886	390	266	117	1	11	0	38
2011	2636	1470	129	44	7	25	1	8
2012	7305	1341	1377	58	7	1	4	1
2013	2268	4801	339	224	4	0	0	1
2014	955	2205	1816	220	77	4	0	1
2015	2163	931	704	232	17	3	0	2
2016	3874	3867	905	573	26	7	1	0
2017	1943	3850	978	69	2	0	0	2

Table S3. Age-specific mass (kg) of North Sea saithe in 1967–2017.

year	age 3	age 4	age 5	age 6	age 7	age 8	age 9	age 10 and older
1967	0.898	1.339	2.094	3.183	3.753	5.316	5.891	7.719
1968	1.234	1.624	1.979	3.007	4.039	4.428	6.136	7.406
1969	0.933	1.530	2.251	2.711	3.558	4.406	5.220	6.767
1970	0.908	1.416	2.049	2.716	3.599	4.463	5.687	6.845
1971	0.811	1.325	2.167	2.934	3.765	4.634	5.172	6.163
1972	0.780	1.175	1.952	2.367	3.793	4.228	4.630	6.326
1973	0.792	1.382	1.633	2.569	3.356	4.684	4.814	6.445
1974	0.831	1.534	2.372	2.751	3.428	4.498	5.713	7.857
1975	0.862	1.472	2.479	3.298	3.764	4.296	5.540	7.562
1976	0.678	1.287	2.250	3.068	4.034	4.383	5.112	7.147
1977	0.733	1.234	1.926	3.108	4.161	4.605	4.859	6.542
1978	0.793	1.304	2.145	3.338	4.521	4.900	5.449	7.400
1979	1.069	1.595	2.228	3.093	4.049	5.274	6.308	7.955
1980	0.921	1.790	2.380	3.028	4.089	5.126	5.939	8.148
1981	0.927	1.790	2.705	3.584	4.535	5.478	6.980	8.724
1982	1.048	1.548	2.518	3.218	4.206	5.125	5.905	8.823
1983	0.992	1.688	2.139	3.135	3.690	4.632	5.505	8.453
1984	0.767	1.586	2.286	2.688	3.895	4.665	6.183	8.474
1985	0.640	1.244	1.941	2.769	3.406	4.950	5.865	8.854
1986	0.670	1.018	1.786	2.430	3.571	4.209	5.651	8.218
1987	0.650	0.861	1.815	3.072	4.209	5.330	6.128	8.603
1988	0.752	0.964	1.379	2.789	4.023	5.254	6.322	8.649
1989	0.864	1.018	1.413	1.997	3.913	5.017	6.430	8.431
1990	0.815	1.175	1.575	2.245	3.241	4.858	6.315	8.416
1991	0.764	1.138	1.744	2.363	3.165	4.222	6.066	8.191
1992	0.930	1.169	1.599	2.240	3.667	4.330	5.412	7.045
1993	0.868	1.239	1.746	2.634	3.184	3.980	5.080	6.891
1994	0.911	1.100	1.594	2.432	3.617	4.787	6.548	8.326
1995	0.967	1.272	1.807	2.560	3.554	4.767	5.267	7.891
1996	0.933	1.167	1.798	2.366	2.951	4.705	6.092	8.382
1997	0.873	1.125	1.445	2.585	3.555	4.525	6.158	8.866
1998	0.861	0.949	1.386	1.743	2.948	3.883	4.996	7.227
1999	0.850	1.042	1.206	1.752	2.337	3.493	4.844	6.745
2000	0.992	1.107	1.532	1.683	2.593	3.084	4.773	7.461
2001	0.774	1.053	1.307	2.093	2.546	3.485	4.141	6.141
2002	0.776	1.014	1.495	1.791	2.961	3.761	4.638	5.750
2003	0.636	0.889	1.167	1.810	2.368	3.176	3.768	5.065
2004	0.794	1.010	1.392	1.896	2.860	3.687	4.814	7.059
2005	0.715	1.155	1.325	1.710	2.132	3.026	3.622	5.713
2006	0.904	1.012	1.489	1.906	2.424	3.058	4.318	5.734
2007	0.769	1.124	1.286	1.834	2.328	2.887	3.600	4.975
2008	0.916	1.065	1.488	1.692	2.210	2.792	3.206	4.565
2009	1.033	1.333	1.672	1.994	2.566	3.086	3.651	4.790
2010	1.037	1.474	2.033	2.597	3.163	3.488	3.968	5.223
2011	0.955	1.192	1.787	2.571	3.068	3.418	3.718	4.289
2012	0.910	1.287	1.383	2.196	3.221	3.536	4.181	4.482
2013	0.878	1.132	1.586	1.957	3.076	3.841	4.541	5.648
2014	1.091	1.265	1.568	2.334	2.607	4.010	5.530	6.679
2015	0.951	1.253	1.621	2.180	3.037	3.793	4.228	7.285
2016	0.937	1.239	1.611	2.231	2.888	3.450	4.331	6.208
2017	0.956	1.228	1.755	2.356	2.987	4.232	4.473	6.287

Table S4. Age-specific mass (kg) of North Sea saithe landings in 1967–2017.

year	age 3	age 4	age 5	age 6	age 7	age 8	age 9	age 10 and older
1967	0.931	1.362	2.104	3.186	3.754	5.316	5.891	7.719
1968	1.278	1.652	1.989	3.009	4.040	4.428	6.136	7.406
1969	0.966	1.557	2.261	2.713	3.559	4.406	5.220	6.768
1970	0.941	1.441	2.059	2.718	3.600	4.463	5.687	6.845
1971	0.840	1.348	2.178	2.936	3.766	4.634	5.173	6.163
1972	0.808	1.196	1.961	2.369	3.794	4.228	4.630	6.326
1973	0.821	1.406	1.641	2.571	3.357	4.684	4.814	6.445
1974	0.861	1.561	2.383	2.753	3.429	4.498	5.713	7.857
1975	0.893	1.498	2.490	3.300	3.765	4.296	5.540	7.562
1976	0.702	1.309	2.260	3.071	4.035	4.383	5.112	7.147
1977	0.760	1.256	1.935	3.111	4.162	4.605	4.859	6.542
1978	0.822	1.327	2.155	3.340	4.522	4.901	5.449	7.400
1979	1.107	1.623	2.238	3.095	4.050	5.274	6.308	7.955
1980	0.955	1.821	2.391	3.030	4.090	5.126	5.939	8.148
1981	0.961	1.821	2.718	3.587	4.536	5.478	6.980	8.724
1982	1.086	1.575	2.529	3.220	4.207	5.125	5.905	8.823
1983	1.028	1.718	2.149	3.138	3.691	4.632	5.505	8.453
1984	0.795	1.614	2.297	2.690	3.896	4.665	6.183	8.474
1985	0.663	1.265	1.951	2.772	3.407	4.950	5.865	8.854
1986	0.694	1.035	1.794	2.432	3.572	4.209	5.651	8.218
1987	0.674	0.876	1.824	3.075	4.210	5.330	6.128	8.603
1988	0.779	0.981	1.386	2.791	4.024	5.254	6.322	8.649
1989	0.895	1.036	1.420	1.998	3.914	5.018	6.430	8.431
1990	0.844	1.196	1.583	2.247	3.242	4.858	6.315	8.416
1991	0.791	1.158	1.752	2.365	3.165	4.222	6.066	8.191
1992	0.964	1.189	1.607	2.242	3.668	4.330	5.413	7.046
1993	0.899	1.260	1.754	2.636	3.185	3.980	5.080	6.891
1994	0.944	1.119	1.601	2.434	3.618	4.787	6.548	8.326
1995	1.002	1.294	1.816	2.562	3.555	4.767	5.267	7.891
1996	0.967	1.187	1.807	2.368	2.952	4.705	6.092	8.382
1997	0.905	1.145	1.452	2.587	3.556	4.525	6.158	8.866
1998	0.892	0.966	1.393	1.744	2.949	3.883	4.996	7.227
1999	0.881	1.061	1.211	1.754	2.337	3.493	4.844	6.745
2000	1.027	1.127	1.539	1.684	2.594	3.084	4.773	7.462
2001	0.802	1.072	1.313	2.095	2.546	3.485	4.141	6.141
2002	0.923	1.035	1.478	1.769	2.947	3.426	4.407	5.674
2003	0.833	0.980	1.173	1.810	2.368	3.176	3.768	5.065
2004	0.918	1.084	1.392	1.896	2.860	3.687	4.814	7.059
2005	0.921	1.155	1.325	1.710	2.132	3.026	3.622	5.713
2006	0.945	1.069	1.514	1.906	2.424	3.058	4.318	5.734
2007	0.837	1.143	1.317	1.840	2.328	2.887	3.600	4.975
2008	0.944	1.193	1.565	1.720	2.226	2.795	3.206	4.565
2009	1.036	1.340	1.664	1.992	2.563	3.085	3.648	4.793
2010	1.036	1.479	2.034	2.597	3.164	3.488	3.968	5.199
2011	1.007	1.207	1.783	2.573	3.068	3.404	3.717	4.284
2012	1.015	1.321	1.408	2.201	3.223	3.536	4.177	4.482
2013	0.898	1.156	1.614	1.976	3.078	3.841	4.541	5.648
2014	1.126	1.300	1.607	2.384	2.617	4.013	5.530	6.679
2015	0.977	1.244	1.625	2.190	3.043	3.796	4.228	7.287
2016	0.998	1.292	1.628	2.283	2.892	3.453	4.333	6.208
2017	1.047	1.302	1.809	2.361	2.988	4.232	4.473	6.292

Table S5. Age-specific mass (kg) of North Sea saithe discards in 1967–2017.

year	age 3	age 4	age 5	age 6	age 7	age 8	age 9	age 10 and older
1967	0.748	1.076	1.818	2.972	3.590	5.316	5.891	7.719
1968	1.028	1.306	1.719	2.808	3.864	4.428	6.136	7.406
1969	0.777	1.230	1.955	2.531	3.403	4.406	5.220	6.767
1970	0.757	1.139	1.780	2.536	3.442	4.463	5.687	6.845
1971	0.676	1.065	1.882	2.739	3.601	4.634	5.172	6.163
1972	0.650	0.945	1.695	2.210	3.628	4.228	4.630	6.326
1973	0.660	1.111	1.419	2.399	3.210	4.684	4.814	6.445
1974	0.692	1.233	2.060	2.568	3.279	4.498	5.713	7.857
1975	0.718	1.184	2.153	3.079	3.600	4.296	5.540	7.562
1976	0.565	1.035	1.954	2.865	3.858	4.383	5.112	7.147
1977	0.611	0.993	1.673	2.902	3.980	4.605	4.859	6.542
1978	0.661	1.049	1.862	3.116	4.325	4.900	5.449	7.400
1979	0.890	1.283	1.935	2.888	3.873	5.274	6.308	7.955
1980	0.768	1.439	2.067	2.827	3.911	5.126	5.939	8.148
1981	0.773	1.439	2.349	3.346	4.338	5.478	6.980	8.724
1982	0.873	1.245	2.186	3.004	4.023	5.125	5.905	8.823
1983	0.826	1.358	1.858	2.927	3.529	4.632	5.505	8.453
1984	0.639	1.276	1.985	2.510	3.726	4.665	6.183	8.474
1985	0.533	1.000	1.686	2.586	3.258	4.950	5.865	8.854
1986	0.558	0.818	1.551	2.269	3.416	4.209	5.651	8.218
1987	0.542	0.693	1.576	2.869	4.026	5.330	6.128	8.603
1988	0.626	0.775	1.198	2.604	3.848	5.254	6.322	8.649
1989	0.720	0.819	1.227	1.865	3.743	5.017	6.430	8.431
1990	0.679	0.945	1.368	2.097	3.100	4.858	6.315	8.416
1991	0.636	0.915	1.515	2.206	3.027	4.222	6.066	8.191
1992	0.775	0.940	1.389	2.092	3.508	4.330	5.412	7.045
1993	0.723	0.996	1.517	2.460	3.046	3.980	5.080	6.891
1994	0.759	0.884	1.384	2.271	3.459	4.787	6.548	8.326
1995	0.806	1.023	1.570	2.390	3.400	4.767	5.267	7.891
1996	0.778	0.938	1.562	2.209	2.823	4.705	6.092	8.382
1997	0.728	0.905	1.255	2.413	3.400	4.525	6.158	8.866
1998	0.717	0.764	1.204	1.627	2.820	3.883	4.996	7.227
1999	0.708	0.838	1.047	1.636	2.235	3.493	4.844	6.745
2000	0.826	0.890	1.330	1.571	2.480	3.084	4.773	7.461
2001	0.645	0.847	1.135	1.955	2.435	3.485	4.141	6.141
2002	0.616	0.896	1.580	2.483	3.469	6.058	6.935	6.927
2003	0.469	0.571	0.641	1.689	2.265	3.176	3.768	5.065
2004	0.617	0.676	1.203	1.769	2.735	3.687	4.814	7.059
2005	0.741	0.913	1.146	1.595	2.038	3.026	3.622	5.713
2006	0.808	0.724	0.859	1.778	2.318	3.058	4.318	5.734
2007	0.660	1.062	0.949	1.365	2.227	2.887	3.600	4.975
2008	0.633	0.680	0.967	1.161	1.495	1.820	3.206	2.797
2009	1.010	1.253	1.946	2.403	2.838	3.388	3.934	3.911
2010	1.046	1.374	1.987	2.561	3.025	3.351	3.968	6.895
2011	0.756	0.971	2.054	2.445	3.170	4.072	4.369	6.618
2012	0.808	0.997	1.101	1.831	2.675	3.411	4.804	5.313
2013	0.835	1.003	1.180	1.300	2.298	3.841	4.541	5.861
2014	0.977	1.072	1.274	1.487	2.077	3.223	5.530	7.568
2015	0.877	1.326	1.531	1.848	2.410	2.184	4.228	5.911
2016	0.882	1.096	1.440	1.764	2.384	2.864	2.634	4.282
2017	0.815	0.937	1.269	1.907	2.484	4.232	4.473	2.817

Table S6. Age-specific maturity (proportion of aduts) of North Sea saithe.

	age 1	age 2	age 3	age 4	age 5	age 6	age 7	age 8 and older
Maturity	0.00	0.00	0.00	0.20	0.65	0.84	0.97	1.00

Table S7. Aggregate exploitable biomass index (catch per unit effort or CPUE) of North Sea saithe in 2000–2017.

year	CUPE
2000	2.104
2001	2.356
2002	1.926
2003	1.794
2004	2.275
2005	2.491
2006	2.584
2007	2.148
2008	2.583
2009	2.026
2010	1.888
2011	1.871
2012	1.641
2013	1.800
2014	1.751
2015	1.973
2016	1.725
2017	1.970

Table S8. Age-specific abundunce indices of North Sea saithe in 1992–2017.

year	age 3	age 4	age 5	age 6	age 7	age 8
1992	1.077	2.760	0.516	0.098	0.057	0.050
1993	7.965	2.781	1.129	0.197	0.011	0.040
1994	1.117	1.615	0.893	0.609	0.091	0.040
1995	13.959	2.501	1.559	0.533	0.172	0.049
1996	3.825	6.533	1.112	0.971	0.212	0.069
1997	3.756	3.351	7.461	0.698	0.534	0.181
1998	1.181	4.134	1.351	1.580	0.149	0.179
1999	2.086	1.907	3.155	0.619	0.632	0.074
2000	3.479	8.836	1.081	0.868	0.114	0.152
2001	21.475	6.169	3.936	0.356	0.444	0.113
2002	10.748	18.974	1.327	1.090	0.162	0.264
2003	19.272	23.802	13.402	0.393	0.439	0.168
2004	4.930	6.727	3.237	0.921	0.064	0.085
2005	8.916	7.512	4.428	1.914	1.082	0.104
2006	10.553	29.579	2.835	1.177	0.445	0.242
2007	34.006	5.578	11.700	1.016	0.743	0.358
2008	3.312	5.584	0.907	1.997	0.254	0.254
2009	1.346	1.703	0.568	0.101	0.229	0.200
2010	1.361	0.964	0.471	0.205	0.045	0.166
2011	4.520	8.451	1.059	1.114	0.426	0.080
2012	11.134	2.497	2.968	0.503	0.483	0.344
2013	14.701	16.279	1.830	1.858	0.308	0.146
2014	1.649	3.923	2.822	0.481	0.520	0.114
2015	11.001	5.613	4.611	1.581	0.289	0.285
2016	37.901	17.439	3.255	2.681	0.945	0.195
2017	11.447	13.102	3.068	1.267	0.942	0.473

Appendix S2. Details of model fitting for the North Sea saithe popuation operating model used in management strategy evaluation.

Figure S1. Process and sampling errors in the North Sea saithe population and observation operating models. (a) recruitment and survival process error and (b) fishing mortality process error estimated from normalized residuals from a state-space estimation model in the 2018 assessment. Standardized residuals for total catch (c), monitoring survey (IBTS–Q3, d), and exploitable biomass index (e) from the 2018 assessment. Blue circles indicate a positive residual and red circles a negative residual.

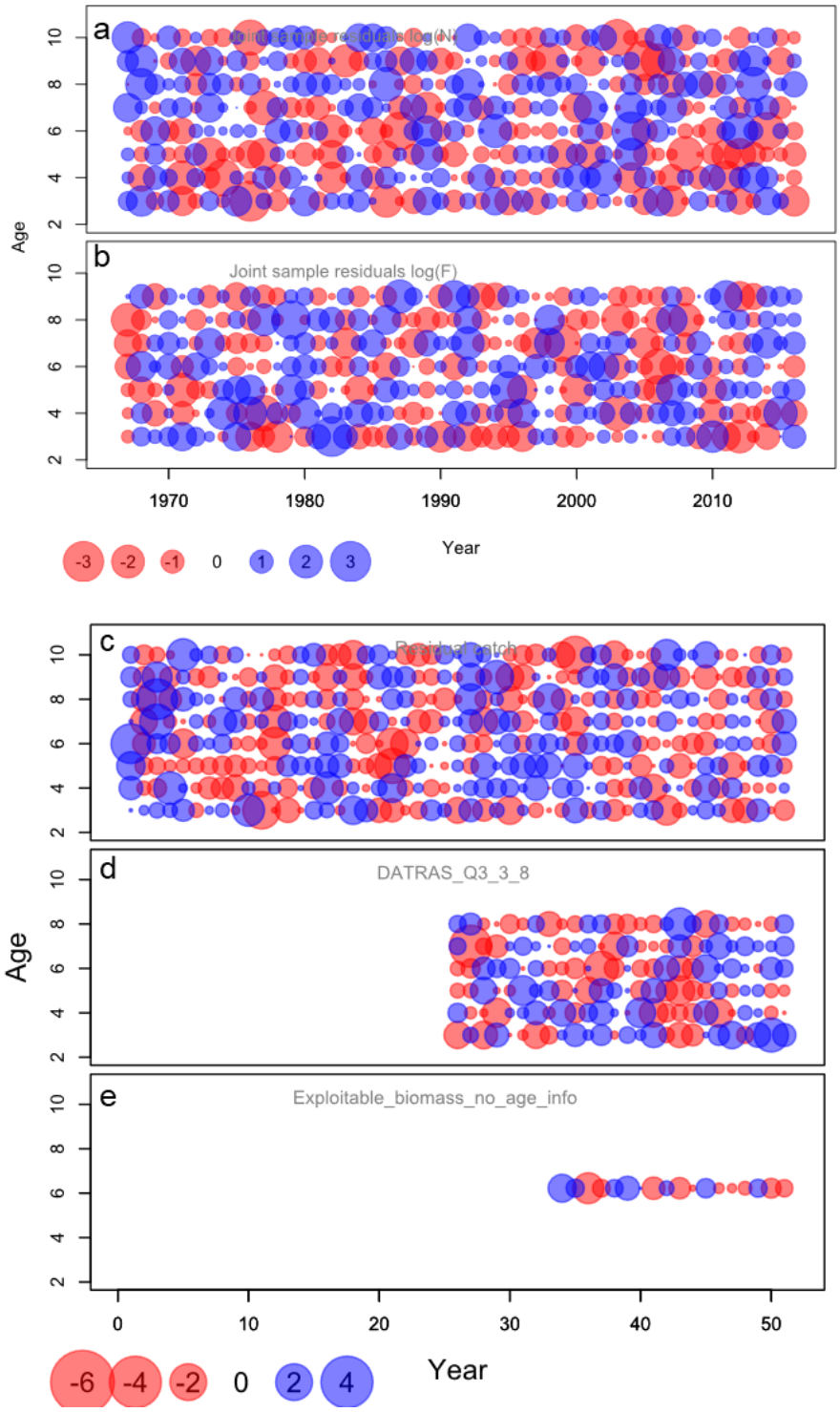

Figure S2. Recruitment in the North Sea saithe operating model. (a) spawner-recruit pairs from historical data (red circles) and those generated from the fitted stock-recruit model (black circles) for 100 replicates. (b) historical (red line) and simulated recruits (black line) using empirical cumulative distribution function (ecdf in R) for the stock-recruit pairs in a. (c) autocorrelation in the 1998–2017 recruit estimates from the 2018 assessment.

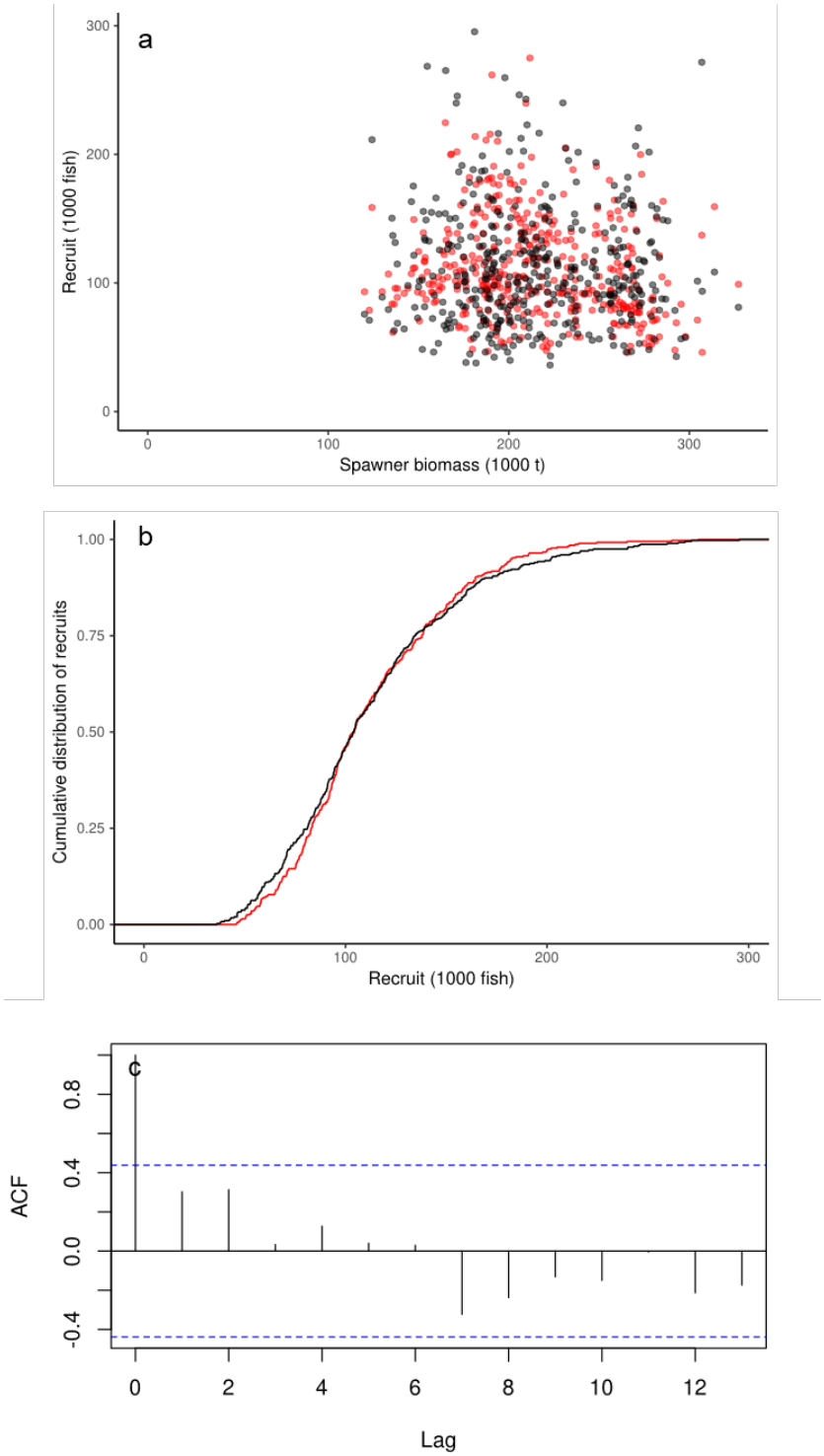
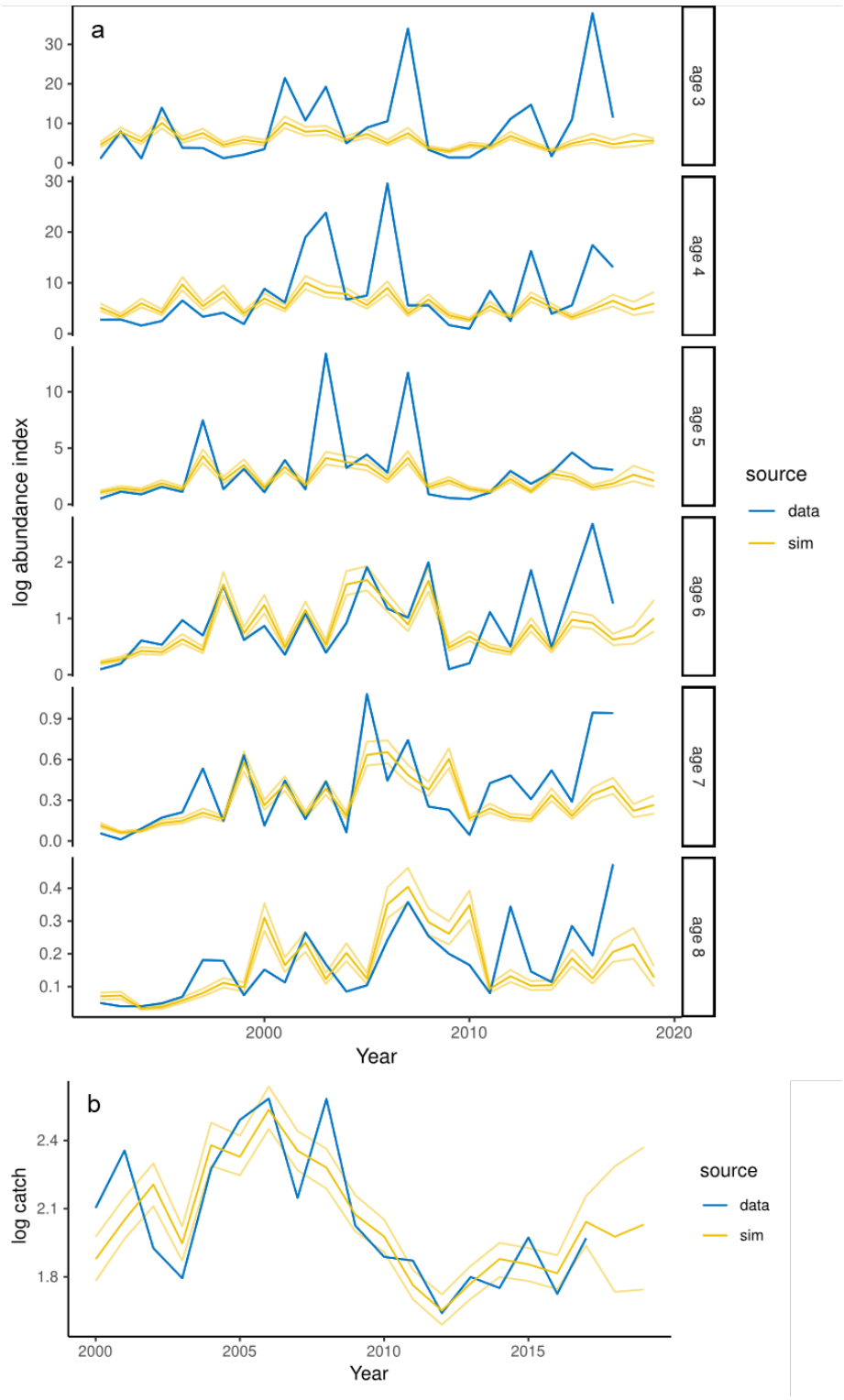
Figure S3. Monitoring and catch surveys in the North Sea saithe operating model. (a) observed (IBTS–Q3) and simulated age-specific number survey indices in 1992–2017. (b) observed and simulated aggregated exploitable biomass index in 2000–2017. For simulated indices, a thick line indicates median, thin lines indicate 75% confidence limits.