A precautionary solution to estimation bias in shaping safe harvest boundaries

Daisuke Goto[1,*], Jennifer A. Devine[1,4], Ibrahim Umar[1], Simon H. Fischer[2], José A. A. De Oliveira[2], Daniel Howell[1], Ernesto Jardim[3,5], Iago Mosqueira[3,6], and Kotaro Ono[1].

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

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

[3]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:

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

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

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

[*]Corresponding author: daisuke.goto2@gmail.com

Abstract

1. Imperfect knowledge of social–ecological systems can obscure predictability of resource fluctuation and, in turn, lead to erroneous risk assessments and delayed management actions. Systematic error in population status such as persistent overestimation of abundance is a pervasive conservation problem and has plagued assessments of commercial exploitation of marine species, threatening its sustainability.

2. Using North Sea saithe (*Pollachius virens*)–a demersal (bottom-water) predatory fish–as a real-world case study, we illustrate a precautionary approach to diagnose robustness of harvest rules to persistent estimation bias (overestimated stock abundance and underestimated fishing mortality rate) and to develop alternative protective measures that minimize population depletion (quasi-extinction) risk by propagating known sources of uncertainty (process, observation, and implementation) through closed-loop simulation of resource–management feedback systems (management strategy evaluation, MSE).

3. Analyses showed that the harvest rules set for saithe are robust to a moderate amount (10–30%) of estimation bias. More severe bias sets overly optimistic catch limits and promotes overexploitation only in the short term; unacceptably high quasi-extinction risks, however, result primarily from progressively amplified amplitudes of catch fluctuation. Although these undesirable outcomes were, to some extent, mitigated by applying a policy tool to suppress catch fluctuation, this tool falls short of being an effective measure to achieve management goals.

4. More consistent performance of management measures was achieved by developing and applying more precautionary harvest rules through MSE by explicitly accounting for bias. When bias became more severe, raising threshold abundance (by 8–24%) that triggers management actions and lowering target exploitation rate (by 6–29%) would not only safeguard against

overexploitation and depletion but also provide catch stability (less disruption in fishing operations).

5. We show that the precautionary approach to risk management through MSE offers a powerful tool to set safe harvest boundaries when assessments are persistently biased. Given challenges in identifying the sources, we suggest bias be routinely evaluated through MSE, and alternative measures be developed to set catch limits when needed. By explicitly accounting for key sources of uncertainty in managing commercial exploitation, our proposed approach ensures effective conservation and sustainable exploitation of living marine resources even under profound uncertainty.

Key words: *environmental stochasticity, decision making, measurement error, risk analysis, extinction, management procedure, state-space models*

1. INTRODUCTION

Managers and policymakers 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). Imperfect knowledge of social–ecological systems, however, 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 (promoting noncompliance and inertia) may also reduce the efficacy of management measures applied (Hilborn et al. 2001; Beddington et al. 2007). We must, thus, account for key sources of uncertainty to accurately assess overexploitation risk (Regan et al. 2005) and recovery potential (Memarzadeh et al. 2019) and set conservation priorities if we were to achieve internationally agreed targets such as Sustainable Development Goal 14 (UN 2015) and Aichi Biodiversity Target 6 (CBD 2010).

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 taking precautionary measures, where necessary. This approach may include the use of biological thresholds such as the population abundance that produces maximum sustainable yield (B_{MSY}, Beddington et al. 2007). The harvest of wild populations is commonly managed by applying decision rules based on such predefined thresholds (reference points) to set a catch limit for the year (Beddington et al. 2007). Accurate population assessments, thus, contribute to successful implementation of management measures to sustain long-term commercial exploitation of marine animals (Hilborn et al. 2020). Scientific uncertainty in assessments (data and models), however, has posed a multitude of challenges (Patterson et al. 2001; Sethi 2010). If population abundance is 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 (resulting heightened 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 (ICES 2020a) and likely contributed to overharvest and depletion including stocks that are considered well-managed (Brooks & Legault 2016; Wiedenmann & Jensen 2018). Inconsistency across assessments (such as persistent overestimation of abundance) detected retrospectively (known as “retrospective patterns”) has led to the rejection of assessments (Punt et al. 2020). Although past research has proposed solutions to the retrospective problems, applying these solutions remains a challenge because bias could originate from multiple sources (Mohn 1999; Hurtado-Ferro et al. 2015; Brooks & Legault 2016). Incomplete knowledge of the causes behind retrospective patterns a priori may lead to incorrect application of the tools, inadvertently exacerbating the problems by amplifying overharvest and depletion risks (Mohn 1999; Brooks & Legault 2016). 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 (Brooks & Legault 2016; Punt et al. 2020).

Here, we illustrate how closed-loop simulation of resource–management systems (management strategy evaluation) can help prevent retrospective patterns from derailing effective management of exploited marine populations under known sources of uncertainty. Management strategy evaluation (MSE) is a flexible decision-support tool frequently used in fisheries (Butterworth & Punt 1999) and has increasingly been adopted for conservation planning of imperiled species in marine and terrestrial systems (Milner-Gulland et al. 2001; Bunnefeld et al. 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 (Punt et al. 2016). MSE can also assess consequences of likely sources of bias in assessments (Szuwalski et al. 2017; Hordyk et al. 2019). Managing with persistent overestimates of population abundance may not only set overly optimistic catch limits (Hordyk et al. 2019) but also amplify the magnitude of catch fluctuation (Deroba 2014), an undesirable outcome for harvesters and seafood processors. In this study, we make use of the MSE framework for the North Sea population of saithe (*Pollachius virens*) (ICES 2019a), a demersal (bottom-water) predatory fish harvested commercially by more than a dozen European nations, as a real-world case study. We illustrate a 2-step simulation approach to diagnose retrospective problems and to design robust harvest policies by explicitly accounting for process, observation, and implementation errors under scenarios of estimation bias (inaccuracy and inconsistency in perceived stock abundance and fishing pressure). Specifically, we ask: 1) How robust are current management procedures to biased assessments? and 2) How precautionary do management procedures need to be to avert mismanagement?

2. METHODS AND MATERIALS

2.1. 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 MSE. We used the MSE framework originally developed for commercially harvested demersal fish stocks in the North Sea through the International Council for the Exploration of the Sea (ICES 2019a) and has been adopted for other managed species in the North Atlantic including Atlantic mackerel (*Scomber scombrus,* ICES 2020b). The framework consists of submodels that simulate 1) true population and harvest dynamics at sea and observations through monitoring surveys (an operating model, OM), and 2) management processes (learning and decision), assessments based on observations from the surveys and subsequent decision making (a management procedure, MP) (Fig. 1a). We conditioned the OM on the latest (2018) assessment for North Sea saithe (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, and projected 21-year (2018–2038) forecasts. We ran all simulations in R (version 3.60, R Development Core Team 2019) using the mse R package (https://github.com/flr/mse) (ICES 2019a), part of the Fisheries Library in R (FLR, Kell et. al 2007).

2.2. Saithe population dynamics

To simulate future saithe population dynamics, we used an age-structured population model that accounts for environmental stochasticity. The data sources, survey methods, and model structure have been extensively documented in ICES (2016) and ICES (2019b). Briefly, we parameterized the model with 51-year estimates of age-specific masses (g) and maturity rates (proportion of adults), and natural mortality rates (non-fishing such as starvation and diseases) 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, t) and age-specific (ages 3–8) abundance indices (International bottom trawl surveys in the third quarter, IBTS-Q3, in 1992–2017) (ICES 2018).

Modeled fish enter the population as 3-year-olds (recruits). We simulated density-dependent regulation of recruitment with a segmented regression (ICES 2019a) relating adult biomass to the number of recruits. 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. 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 S1a,b); this process was repeated independently for each replicate. Preliminary analyses showed little evidence of temporal autocorrelation in recruitment (Appendix S1c).

We simulated the population dynamics of 4-year-olds and older as

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

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

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

where $N_{a,y}$, $F_{a,y}$, and $M_{a,y}$ are a-year-old numbers, fishing mortality rates, and natural mortality 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). 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 S2a, ICES 2019b). 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.

2.3. 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 (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 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; 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 S2b).

We simulated uncertainty in reported catch by computing a commercial catch index (or exploitable biomass index) generated from the population OM (Appendix 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 S2c).

2.4. 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) before provision of catch advice in May. In this study, we used the State-space Assessment Model (Nielsen & Berg 2014) as an EM and harvest rules set for saithe (ICES 2019a); model settings and forecast assumptions are fully described in ICES (2019b). 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 2 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).

2.5. Population and management measure performances

We computed conservation-oriented (median SSB and risk) and harvest-oriented (median catch and interannual catch variability, ICV) metrics from 1000 realizations of annual assessments to evaluate performance of the harvest rules applied. We chose the number of replicates based on the stability of risk (ICES 2019a). Risk is defined as the maximum annual probability of SSB falling below a limit threshold, B_{lim} (probability of quasi-extinction, Fig. 1b), consistent with previous analyses (ICES 2019c). B_{lim} is a spawner abundance below which reproductive capacity of the population is expected to decline (Rindorf et al. 2016). 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 2019a). We computed ICV (a percentage change in catch limits) as

$$\mathrm{ICV}_y = \frac{C_{y+1} - C_y}{C_y} \times 100 \qquad (4)$$

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

We computed Mohn's ρ, which indicates the degree of inconsistency (bias) between subsequent assessments ("retrospective pattern", Mohn 1999), and relative error (proportional deviation between the population OM and EM) for SSB and fishing mortality rate (mean F, which is computed from age-specific fishing mortality rates of 4- to 7- year-olds) to evaluate the performance of annual assessments. We computed ρ as mean relative bias with a 5-year moving window ("peel") for the forecasting period in 2027–2038 as;

$$\rho = \frac{1}{5}\sum_{i=1}^{5}\left(\frac{\hat{\theta}_{T-i}^{R_i} - \hat{\theta}_{T-i}}{\hat{\theta}_{T-i}}\right) \quad (5)$$

where $\hat{\theta}_{T-i}^{R_i}$ is an estimate of SSB or mean *F* in the terminal year from the EM, *T*, with the last *i* years of data removed ("peeled"), and $\hat{\theta}_{T-i}$ is the estimate for year *T-i*, with all data included (Brooks & Legault 2016).

2.6. Estimation bias scenarios

To evaluate how managing with persistent estimation bias degrades performance of harvest rules and, in turn, potential to achieve management goals, we simulated 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. In this study, we analyzed 11 scenarios that can cause severe conservation issues for exploited species, SSB overestimation and mean *F* underestimation. Specifically, we added a positive bias (+0%/baseline, +10%, +20%, +30, +40%, and +50% per year) to age-specific numbers or a negative bias (–0%/baseline, –10%, –20%, –30, –40%, and –50% per year) to age-specific fishing mortality rates in the terminal year of annual assessments.

We considered three harvest rules evaluated for saithe in previous analyses (Fig. 1b, ICES 2019c); 2 rules (HCR-A+D and HCR-A1+D) with harvest policies of interannual catch quota flexibility and suppressing short-term catch fluctuation (hereafter stability constraint) and 1 rule without (HCR-A). For simplicity, we introduced interannual catch quota flexibility (also known as "banking and borrowing", ICES 2019a) by simulating a scenario of over- and under-harvesting by a fixed proportion (10%) in alternate years (for example, 10% of catch quota in year 1 is transferred to catch quota in year 2, and so on). In effect, this scenario can act like implementation error. Stability constraints are designed to suppress year-to-year variability in projected catch to ≤ 25% upward and to ≤ 20% downward (moderate) under HCR-A+D, or to ≤ 15% in both directions (strict) under HCR-A1+D. By running these scenarios, we evaluated how effective the policy tool designed to provide stable and predictable catch forecasts is when the assessments are biased. For consistency, we used the same decision model parameter values as in ICES (2019a) in all analyses (HCR-A: $B_{trigger}$ = 250,000 t and F_{target} = 0.35; HCR-A+D: $B_{trigger}$ = 230,000 t and F_{target} = 0.37; HCR-A1+D: $B_{trigger}$ = 230,000 t and F_{target} = 0.36, Fig. 1b, ICES 2019a). We analyzed all scenarios based on the performance metrics (median SSB, risk, median catch, and ICV) from 1000 realizations of short-term (years 1–5) and long-term (years 11–20) projections.

2.7. Developing robust management measures

To evaluate how precautionary the harvest rules need to be to minimize disruption in catch advice provisioning when the assessment is biased, we explored alternative rules by optimizing the 2 parameters of the decision model ($B_{trigger}$ and F_{target}) to project catch limits under the same bias scenarios (overestimated SSB or underestimated mean F by 10–50%) through MSE. Because this is a computationally intensive procedure, we explored select candidate combinations (192 per scenario or 1920 unique runs in total) using HCR-A 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 by maximizing median catch limits while maintaining long-term risk ≤ 0.05 (ICES 2019c).

3. RESULTS

3.1. Estimation errors under the baseline scenario

Analyses on the baseline scenario forecasts showed that process and observation errors introduced minor inaccuracy (relative error = ~–0.02%) and inconsistency (Mohn's ρ = ~–0.002) in annual estimates of SSB and mean *F* with the EM under all the harvest rules (Table 1). Implementation error with stability constraints amplified inconsistency but did not influence inaccuracy except for mean *F* under the strict constraint in which the EM failed to detect an increase (Table 1). Mean ρ for SSB or mean *F* did not correlate strongly with relative errors (correlation coefficient r = ~0.01 and ~0.008, respectively) under any of the harvest rules.

3.2. Performance of harvest rules with biased estimates

Overall, all harvest rules were more robust to bias in estimated mean *F* than SSB. An increasing amount (10% to 50%) of bias in the estimates was projected to increase median catch, reduce SSB, and, in turn, increase risk in the short term (Fig. 2). Without accounting for implementation error or constraining year-to-year catch variability (HCR-A), bias in SSB estimates led to as much as 18% more catch, 14% less SSB, and nearly 3.0x higher risk than bias in mean *F* estimates (Fig. 2a). Although mean ICV moderately increased with bias in SSB estimates, the distribution was highly skewed; by contrast, ICV declined with increasing bias in mean *F* estimates (Fig. 2a).

When accounting for implementation error and applying stability constraints (HCR-A+D and HCR-A1+D), similar short-term patterns emerged, except that catch increased less with biased SSB estimates (Fig. 2b,c). By contrast, ICV increased much more with biased estimates of SSB and mean *F* (up to 3.8x and 44.0x, respectively), further increasing risk (up to 1.6x and 2.8x, respectively), under the moderate constraint than under no constraint (Fig. 2b). Under the strict constraint, the distribution of ICV became less skewed, and mean ICV became less responsive to increasing bias, lowering risk by as much as ~15% (Fig. 2c).

In the long-term, although median catch became less responsive (declined by less than 5%) to increasing bias in estimated SSB and mean *F*, ICV became more variable (as much as ~1.4x the short term), reducing SSB and, in turn, increasing risk 12.9- and 8.9- fold under HCR-A (respectively, Fig. 3a). To quantify more precisely how robust the harvest rule is to bias, we ran additional simulations with 1% increments (+10% to +20% for SSB and –20% to –30% for mean *F*). The harvest rule was not precautionary when more than +18% bias in SSB estimates and –24% bias in mean *F* estimates were introduced (Appendix S4).

With implementation error and stability constraints, long-term relative responses of median catch and SSB to increasing bias in SSB and mean *F* estimates were similar to those under HCR-A, but ICV and risk increased less (Fig. 3b,c). Although absolute values of ICV and risk were higher under HCR-A+D and HCR-A1+D owing to implementation error, stability constraints reduced relative changes in ICV and risk (by as much as 26% and 59%, respectively) with increasing bias (especially in mean *F* estimates, Fig. 3b,c).

3.3. Alternative management measures

The proportion of the select grid search area evaluated through MSE that remained precautionary (safe harvest margin) progressively shrank as more bias in SSB and mean *F* estimates was introduced (from 84% to 29% and from 90% to 37%, respectively, Fig. 4a,b and Table 2). Within the safe harvest margin, the harvest rule was projected to produce higher (by 6.7–25%) short-term catches and maintain similar (<3% deviation from the baseline) long-term catches under all bias scenarios (Table 2). With overestimated stock abundances, the fishery produced highest catch limits at lower (by 0.02–0.10) F_{target} and higher (by 10,000–60,000 t) $B_{trigger}$ (Table 2 and Fig. 4a). However, short- and long-term SSB and short-term ICV progressively declined with biased estimates (Table 2). Similarly, with underestimated mean F, the fishery produced higher catches and reduced short-term ICV with lower SSB at lower F_{target} and $B_{trigger}$ (by 0.02–0.06 and 20,000 t, respectively, Table 2 and Fig. 4b).

4. DISCUSSION

We showed that a precautionary approach applied through MSE offers a powerful decision-support tool to explicitly evaluate how robust harvest rules are to estimation bias and, when necessary, to develop alternative (reliable) measures for sustainable harvest of marine populations by simulating the entire commercial fishery system. For North Sea saithe, the current harvest rule is robust to a moderate amount (10–30%) of bias in assessments despite process, observation, and implementation uncertainties. More severe bias sets overly optimistic catch limits only in the short term; unacceptably high risks of missing management targets (quasi-extinction), however, result primarily from progressively amplified fluctuation in catch limits over time. A harvest policy tool to suppress catch fluctuation can, to some extent, mitigate these undesirable effects. More consistent, cost-effective performance–lower risk with less disruption in fishing operations (more stable catch limits)–can be achieved by developing and applying more precautionary harvest rules with lower target exploitation rate and higher threshold abundance. By explicitly accounting for key sources of uncertainty in managing commercial exploitation, this approach can provide decision makers a means to balance common trade-offs–achieving socioeconomic goals while conserving living marine resources.

How robust management measures are to estimation errors would depend on life history, fishing operations, and current status of a given species or population (Wiedenmann & Jensen 2018). For depleted populations with a high growth rate, for example, even a modest amount of bias may incur ecological and socioeconomic damages under multiple sources of uncertainty. North Sea saithe is currently (2018) in good condition (~37% above MSY $B_{trigger}$, ICES 2019b); our MSE-guided analyses showed that the current harvest rule is robust up to 18% overestimation of spawner abundance and 24% underestimation of fishing pressure (when accounting for process and observation errors), which are roughly in agreement with the rule of thumb based on Mohn's ρ proposed by Hurtado-Ferro et al. (2015). Past work, however, suggests that this metric may not be consistently sensitive to retrospective patterns (Hurtado-Ferro et al. 2015; Brooks & Legault 2016; Wiedenmann & Jensen 2018). Our analyses also showed that Mohn's ρ does not reflect the magnitude of relative error (simulated-true versus perceived) in assessments, leaving certain ambiguity in its use.

Our simulations further revealed that managing harvest with severely biased assessments can increase the risk of 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 catch limits. Although median catch limits eventually stabilize, year-to-year catch variance continues to rise (by as much as 74%) over time as the estimates of stock abundance and fishing pressure become progressively more biased, and thus, the risk remains elevated.

Ignorance of retrospective patterns can have time-varying consequences for managers and stakeholders; decision making misguided by erroneous assessments would produce higher yield (and thus revenues) in the short term but ultimately would amplify catch fluctuation and probabilities of depletion and quasi-extinction (and, in turn, fishery closure) in the long term.

Trade-offs between short-term gains and long-term losses are common dilemmas in managing natural resources (Carpenter et al. 2015). Although in-depth analyses on management measures to achieve a balance between conservation and socioeconomic targets are beyond the scope of this study, our findings reemphasize that alternative protective measures need to be explicitly assessed before implementation when providing a scientific basis to inform defensible decision making.

Large year-to-year fluctuation in catch forecasts is disfavored by fishing communities (Carpenter et al. 2015); thus, measures to suppress the fluctuation is commonly applied in industrial exploitation. In our saithe example, this policy tool falls short of being an effective measure to achieve conservation- and harvest-oriented goals under severe uncertainty. Although suppressing short-term catch fluctuation can attenuate catch variance inflated by underestimated fishing pressure (but not overestimated stock abundance), quasi-extinction risk remains unacceptably high under most of the bias scenarios tested. Thus, this strategy may not be sufficiently sensitive to rapid population declines under severe bias in assessments and unlikely prompts reductions in catch effectively.

Our analyses suggest that retrospective problems could go unnoticed for a long time as persistent overestimation of abundance can mask overharvesting and depletion, thereby delaying management responses (asynchronized resource–fishery dynamics, Fryxell et al. 2010). Although a certain lag in management responses is unavoidable, severe retrospective patterns can contribute to management inertia. Once population abundance reaches to a lower threshold (B_{lim}, for example), the population may even become unresponsive to any measure for recovery (Allee effect, Kuparinen et al. 2014). Our analyses showed that this undesirable state can be avoided by developing and applying alternative–more precautionary–harvest rules to set catch limits. For saithe, when retrospective patterns become severe, lowering target exploitation rate and raising threshold abundance (that trigger management actions) would not only minimize probabilities of quasi-extinction and fishery closure but also maintain catch stability, thereby minimizing disruption in fishing operations. Thus, this approach would support cost-effective decision making to safeguard against ecologically and socioeconomically undesirable outcomes of managing risks under systematic uncertainty.

We showed that MSE offers a precautionary solution to retrospective problems in assessments and management of exploited populations. MSE can not only act as a diagnostic tool to evaluate the robustness of management measures by explicitly accounting for long-term consequences but also provide an adaptive, transparent approach to develop alternative protective measures when the perception deviates too far from the reality. Given ubiquity of estimation errors and challenges in identifying the sources (Hurtado-Ferro et al. 2015; Brooks & Legault 2016; Szuwalski et al. 2017), we suggest retrospective patterns be routinely evaluated through MSE as an additional source of uncertainty, and alternative measures be developed to set catch limits when the uncertainty becomes too severe.

Our proposed approach also has limitations. Analyses showed that 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”, Anderies et al. 2019) as the magnitude of retrospective patterns increases. Thus, continued efforts to develop methods to identify root causes of the uncertainty (such as temporal variability in life-history traits, Hurtado-Ferro et al. 2015; Szuwalski et al. 2017) are needed.

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).

Furthermore, changing ocean conditions are projected to increase environmental stochasticity, amplifying marine population and harvest fluctuation (Brooks & Legault 2016). Higher environmental stochasticity may also promote autocorrelation in population fluctuation (Ripa & Lundberg 1996; Gamelon et al. 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 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 MSE work. We also thank 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.

AUTHORS' CONTRIBUTIONS

J.A.A.D., D.H., and D.G. conceived the ideas; J.A.A.D., D.H., S.H.F., I.U., E.J., I.M., J.A.D., and D.G. designed methodology; J.A.D. and D.G. analyzed and interpreted the data; D.G. led the writing of the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

REFERENCES

Anderies JM, Mathias J-D, Janssen MA. 2019. Knowledge infrastructure and safe operating spaces in social–ecological systems. Proceedings of the National Academy of Sciences **116**:5277-5284.

Beddington JR, Agnew DJ, Clark CW. 2007. Current problems in the management of marine fisheries. Science **316**:1713-1716.

Brooks EN, Legault CM. 2016. Retrospective forecasting—evaluating performance of stock projections for New England groundfish stocks. Canadian Journal of Fisheries and Aquatic Sciences **73**:935-950.

Bunnefeld N, Hoshino E, Milner-Gulland EJ. 2011. Management strategy evaluation: a powerful tool for conservation? Trends in Ecology & Evolution **26**:441-447.

Butterworth D, Punt A. 1999. Experiences in the evaluation and implementation of management procedures. ICES Journal of Marine Science **56**:985-998.

Carpenter SR, Brock WA, Folke C, Van Nes EH, Scheffer M. 2015. Allowing variance may enlarge the safe operating space for exploited ecosystems. Proceedings of the National Academy of Sciences **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.

Charles AT. 1998. Living with uncertainty in fisheries: analytical methods, management priorities and the Canadian groundfishery experience. Fisheries Research **37**:37-50.

Costello C, Cao L, Gelcich S, Cisneros-Mata MÁ, Free CM, Froehlich HE, Golden CD, Ishimura G, Maier J, Macadam-Somer I. 2020. The future of food from the sea. Nature:1-6.

Deroba JJ. 2014. 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.

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

Fryxell JM, Packer C, McCann K, Solberg EJ, Sæther B-E. 2010. Resource management cycles and the sustainability of harvested wildlife populations. Science **328**:903-906.

Gamelon M, Sandercock BK, Sæther BE. 2019. Does harvesting amplify environmentally induced population fluctuations over time in marine and terrestrial species? Journal of Applied Ecology **56**:2186-2194.

Hilborn R, Amoroso RO, Anderson CM, Baum JK, Branch TA, Costello C, de Moor CL, Faraj A, Hively D, Jensen OP. 2020. Effective fisheries management instrumental in improving fish stock status. Proceedings of the National Academy of Sciences.

Hilborn R, Maguire J-J, Parma AM, Rosenberg AA. 2001. The precautionary approach and risk management: can they increase the probability of successes in fishery management? Canadian Journal of Fisheries and Aquatic Sciences **58**:99-107.

Hordyk AR, Huynh QC, Carruthers TR. 2019. Misspecification in stock assessments: Common uncertainties and asymmetric risks. Fish and Fisheries **20**:888-902.

Hurtado-Ferro F, Szuwalski CS, Valero JL, Anderson SC, Cunningham CJ, Johnson KF, Licandeo R, McGilliard CR, Monnahan CC, Muradian ML. 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.

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

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

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

ICES. 2020a. Workshop on Catch Forecast from Biased Assessments (WKFORBIAS). ICES Scientific Reports. 2:28. 38 pp.

ICES. 2020b. Workshop on Management Strategy Evaluation of Mackerel (WKMSEMAC). ICES Scientific Reports. 2:74. 175 pp.

Kell LT, 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 Journal of Marine Science 64:640-646.

Kuparinen A, Keith DM, Hutchings JA. 2014. Allee effect and the uncertainty of population recovery. Conservation Biology **28**:790-798.

Memarzadeh M, Britten GL, Worm B, Boettiger C. 2019. Rebuilding global fisheries under uncertainty. Proceedings of the National Academy of Sciences **116**:15985-15990.

Milner-Gulland E, Shea K, Possingham H, Coulson T, Wilcox C. 2001. Competing harvesting strategies in a simulated population under uncertainty. Pages 157-167. Animal Conservation forum. Cambridge University Press.

Mohn R. 1999. The retrospective problem in sequential population analysis: an investigation using cod fishery and simulated data. ICES Journal of Marine Science **56**:473-488.

Nielsen A, Berg CW. 2014. Estimation of time-varying selectivity in stock assessments using state-space models. Fisheries Research **158**:96-101.

Ovaskainen O, Meerson B. 2010. Stochastic models of population extinction. Trends in Ecology & Evolution **25**:643-652.

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

Punt AE, Butterworth DS, de Moor CL, De Oliveira JA, Haddon M. 2016. Management strategy evaluation: best practices. Fish and Fisheries **17**:303-334.

Punt AE, Tuck GN, Day J, Canales CM, Cope JM, de Moor CL, De Oliveira JAA, Dickey-Collas M, Elvarsson BÞ, Haltuch MA. 2020. When are model-based stock assessments rejected for use in management and what happens then? Fisheries Research **224**:105465.

Regan HM, Ben-Haim Y, Langford B, Wilson WG, Lundberg P, Andelman SJ, Burgman MA. 2005. Robust decision□making under severe uncertainty for conservation management. Ecological Applications **15**:1471-1477.

Rindorf A, Cardinale M, Shephard S, De Oliveira JA, Hjorleifsson E, Kempf A, Luzenczyk A, Millar C, Miller DC, Needle CL. 2016. Fishing for MSY: can “pretty good yield” ranges be used without impairing recruitment. ICES Journal of Marine Science.

Ripa J, Lundberg P. 1996. Noise colour and the risk of population extinctions. Proceedings of the Royal Society of London. Series B: Biological Sciences **263**:1751-1753.

Sethi SA. 2010. Risk management for fisheries. Fish and Fisheries **11**:341-365.

Szuwalski CS, Ianelli JN, Punt AE. 2017. Reducing retrospective patterns in stock assessment and impacts on management performance. ICES Journal of Marine Science **75**:596-609.

UN. 2015. Transforming our world: the 2030 Agenda for Sustainable Development A/RES/70/1, Available from https://sustainabledevelopment.un.org/post2015/transformingourworld (accessed 26 September 2020).

Walters C, Maguire J-J. 1996. Lessons for stock assessment from the northern cod collapse. Reviews in Fish Biology and Fisheries **6**:125-137.

Wiedenmann J, Jensen OP. 2018. Uncertainty in stock assessment estimates for New England groundfish and its impact on achieving target harvest rates. Canadian Journal of Fisheries and Aquatic Sciences **75**:342-356.

Tables

Table 1. Summary statistics of performance metrics from North Sea saithe management strategy evaluation under the baseline scenario of three harvest rules evaluated in this study (HCR-A, HCR-A+D, and HCR-A1+D).

	OM[a]				EM[a]				relative error				Mohn's ρ			
harvest rule	median	SD	5th	95th	median	SD	5th	95th	median	Sd	5th	95th	median	SD	5th	95th
SSB[a]																
HCR-A	277556	104612	150732	482385	267070	101053	152967	474471	-0.022	0.278	-0.359	0.507	0.007	0.043	-0.066	0.074
HCR-A+D	216907	104455	81966	419176	221793	100485	80162	411952	-0.010	0.467	-0.467	0.844	0.008	0.047	-0.072	0.076
HCR-A1+D	265121	100132	142147	460374	256608	96271	144700	454235	-0.017	0.293	-0.358	0.538	0.008	0.041	-0.065	0.069
mean F[a]																
HCR-A	0.339	0.143	0.155	0.604	0.340	0.130	0.173	0.574	-0.019	0.551	-0.476	1.062	-0.010	0.081	-0.123	0.143
HCR-A+D	0.339	0.143	0.155	0.604	0.340	0.130	0.173	0.574	-0.019	0.551	-0.476	1.062	-0.012	0.080	-0.123	0.138
HCR-A1+D	0.348	0.162	0.152	0.658	0.340	0.145	0.171	0.619	-0.020	0.821	-0.550	1.432	-0.013	0.076	-0.116	0.137

[a]SSB, mean *F*, OM, and EM indicate spawning stock biomass, mean fishing mortality rates of 4- to 7- year-olds, the population operating model, and the estimation model, respectively.

Table 2. Optimized parameters (F_{target} and $B_{trigger}$)[a] of the harvest rule set for North Sea saithe (HCR-A) and performance metrics[b] from management strategy evaluation under scenarios of varying levels of assessment bias.

			short-term				long-term				
scenario[c]	F_{target}	$B_{trigger}$	catch	IAV	SSB	risk[d]	catch	IAV	SSB	risk[d]	SHM[e]
base	0.35	250000	92464	20.4	251973	2.0	116700	17.7	292067	1.5	-
10% *N*	0.33	250000	101786	13.3	238194	3.2	116288	17.8	279135	2.5	84.4
20% *N*	0.31	270000	103545	12.5	235356	3.3	116154	18.7	274958	3.0	65.6
30% *N*	0.27	310000	93047	20.0	252123	2.2	115984	18.0	293711	2.2	53.1
40% *N*	0.26	310000	101131	13.8	240643	2.9	115863	18.4	282929	2.5	37.5
50% *N*	0.25	310000	104943	12.2	234525	3.3	115730	19.1	274228	2.8	29.2
10% *F*	0.35	230000	103441	12.3	234493	3.3	116897	17.5	272497	2.9	89.6
20% *F*	0.33	230000	102882	12.6	235089	3.3	117221	16.9	273309	3.0	76.0
30% *F*	0.32	230000	104840	11.7	230922	3.6	117677	17.1	267727	3.6	62.5
40% *F*	0.30	230000	103894	11.9	232210	3.5	118376	16.5	269327	3.5	51.0
50% *F*	0.29	230000	105721	11.1	227980	4.1	118942	17.0	262836	4.2	36.5

[a]The model parameters were optimized at the highest median catch while risk remains $\leq$ 5%.
[b]The performance was evaluated with short-term (years 1–5) and long-term (years 11–20) median catch (t), interannual catch variability (%, IAV), median spawning stock biomass (SSB, t), and risk (%).
[c]Scenarios simulate overestimation of abundance (*N*) and underestimation of fishing mortality rate (*F*).
[d]Risk is the maximum probability of SSB falling below B_{lim} (107,297 t) over a given period.
[c]Safe harvest margin (SHM) indicates the proportion of the grid-search area with the harvest rules that remain precautionary (Fig. 4).

Figures legends

Figure 1. Management strategy evaluation (MSE) framework and historical population and harvest dynamics of North Sea saithe. (a) schematic of the MSE framework (FLR/a4a, redrawn from E. Jardim, https://github.com/ejardim) adopted for evaluation of saithe management strategies. (b) three harvest control rules (HCR-A, HCR-A+D, and HCR- A1+D) evaluated in this study. (c) reconstructed saithe population and harvest dynamics taken from the 2018 assessment. In b, blue dashed (horizontal and vertical) lines indicate the harvest rule parameters (reference points) set for saithe (F_{target} and $B_{trigger}$, respectively).

Figure 2. Short-term (years 1–5) performance of management strategies (a) HCR-A, (b) HCR-A+D, and (c) HCR-A1+D for North Sea saithe under scenarios (11) 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 3. Long-term (years 11–20) performance of management strategies (a) HCR-A, (b) HCR-A+D, and (c) HCR-A1+D for North Sea saithe under scenarios (11) of varying levels of estimation bias (overestimation of stock abundance and underestimation of fishing mortality rate). See Fig. 2 for more details.

Figure 4. Grid search for combination of the decision model (harvest rule, HCR-A) parameters (F_{target} and $B_{trigger}$) for North Sea saithe under scenarios (11) 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 $\leq$ 5%) in the long term (years 11–20). Black rectangles indicate combinations of the model parameters with the highest median catch.

Figures

Figure 1.

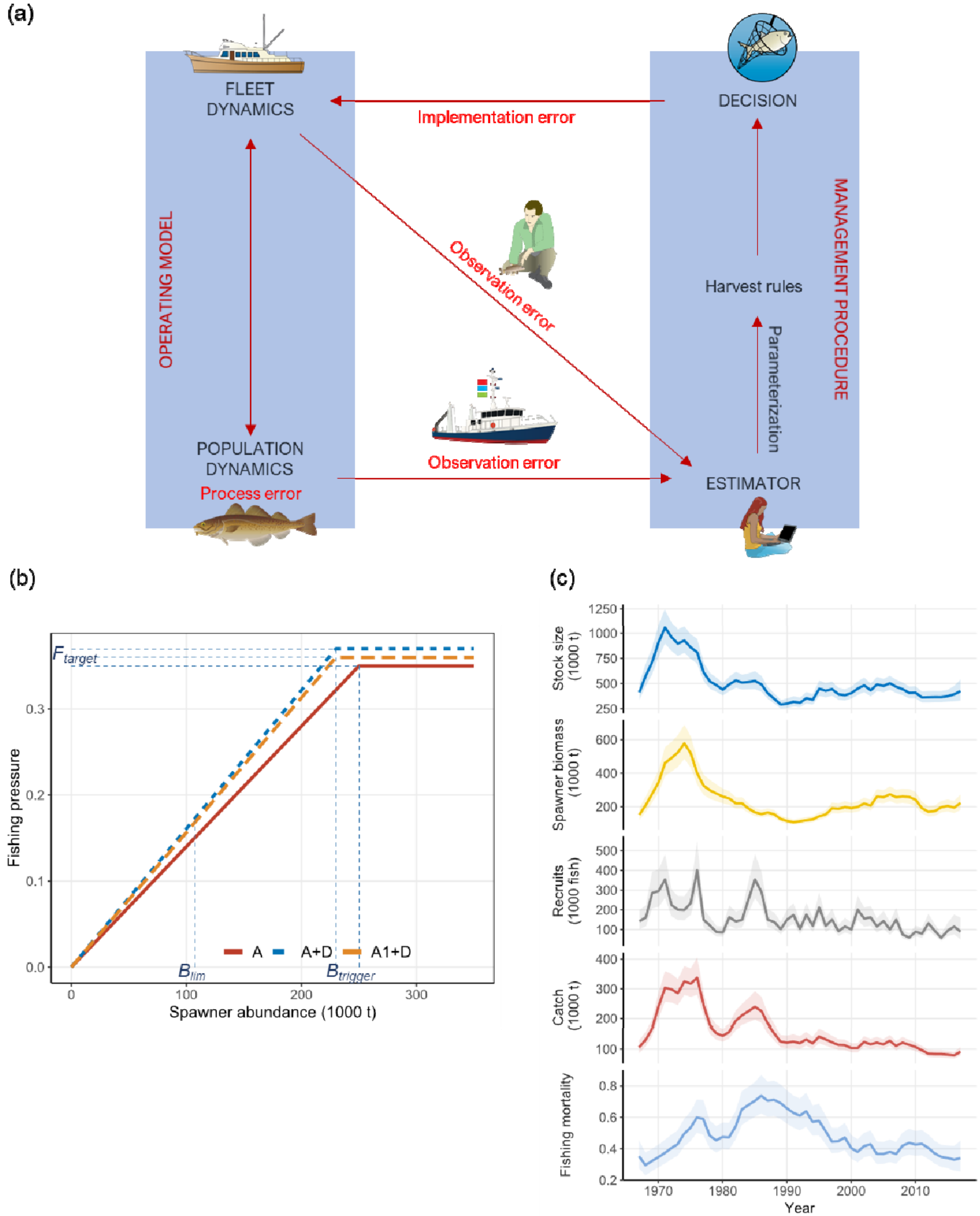


Figure 2.

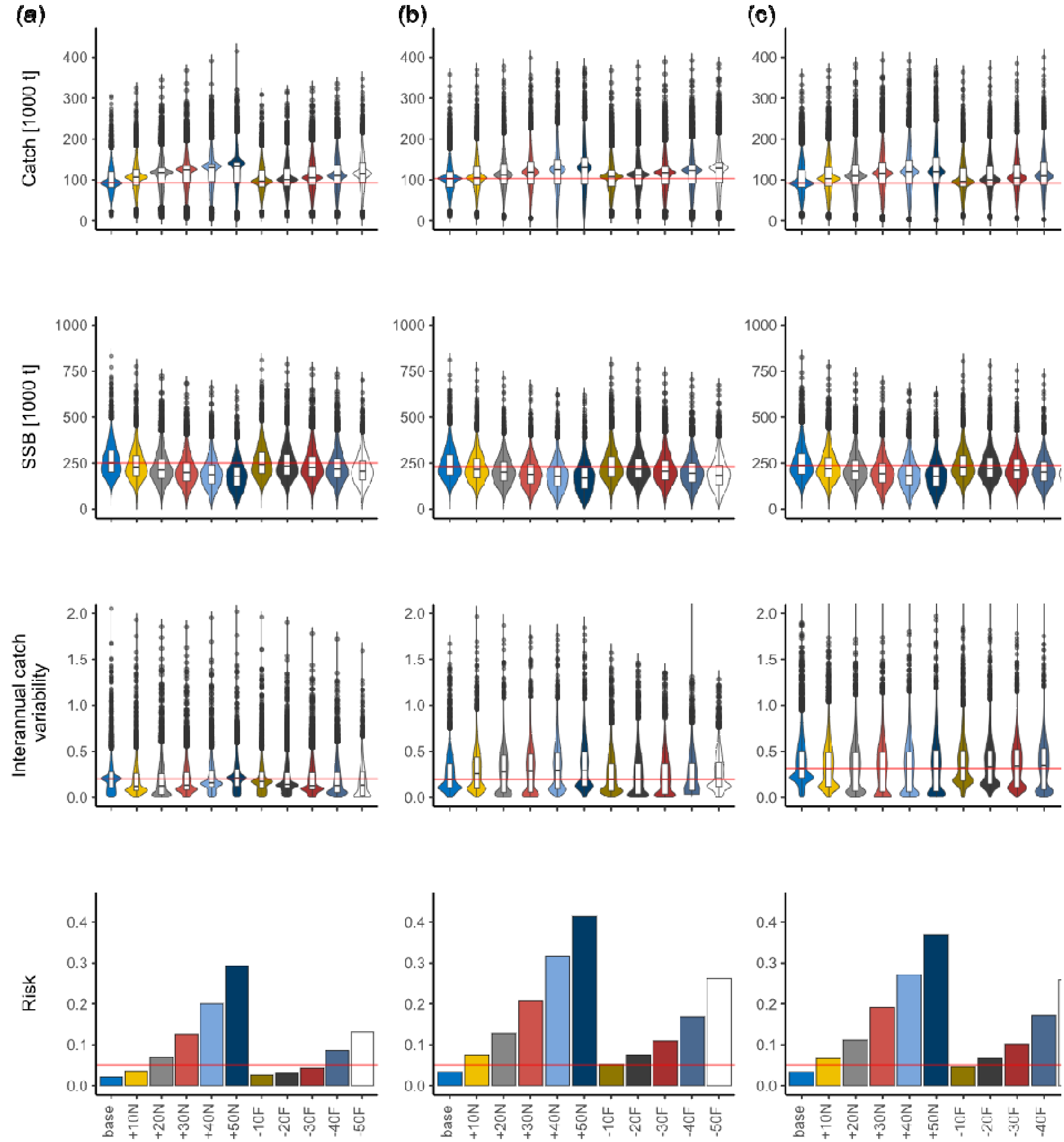

(a)
(b)
(c)
Catch [1000 t]
SSB [1000 t]
Interannual catch variability
Risk
base
+10N
+20N
+30N
+40N
+50N
-10F
-20F
-30F
-40F
-50F

Figure 3.

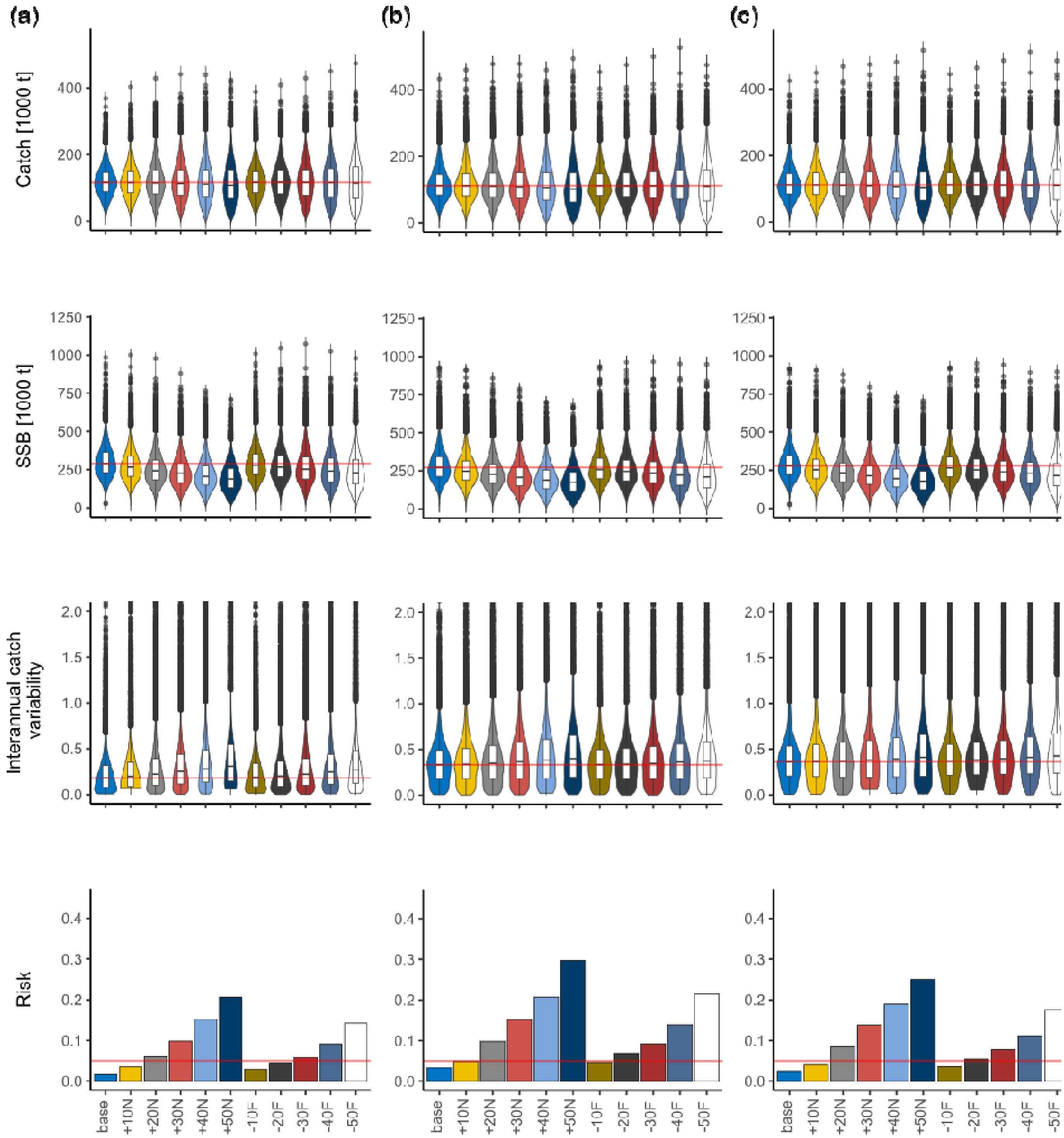
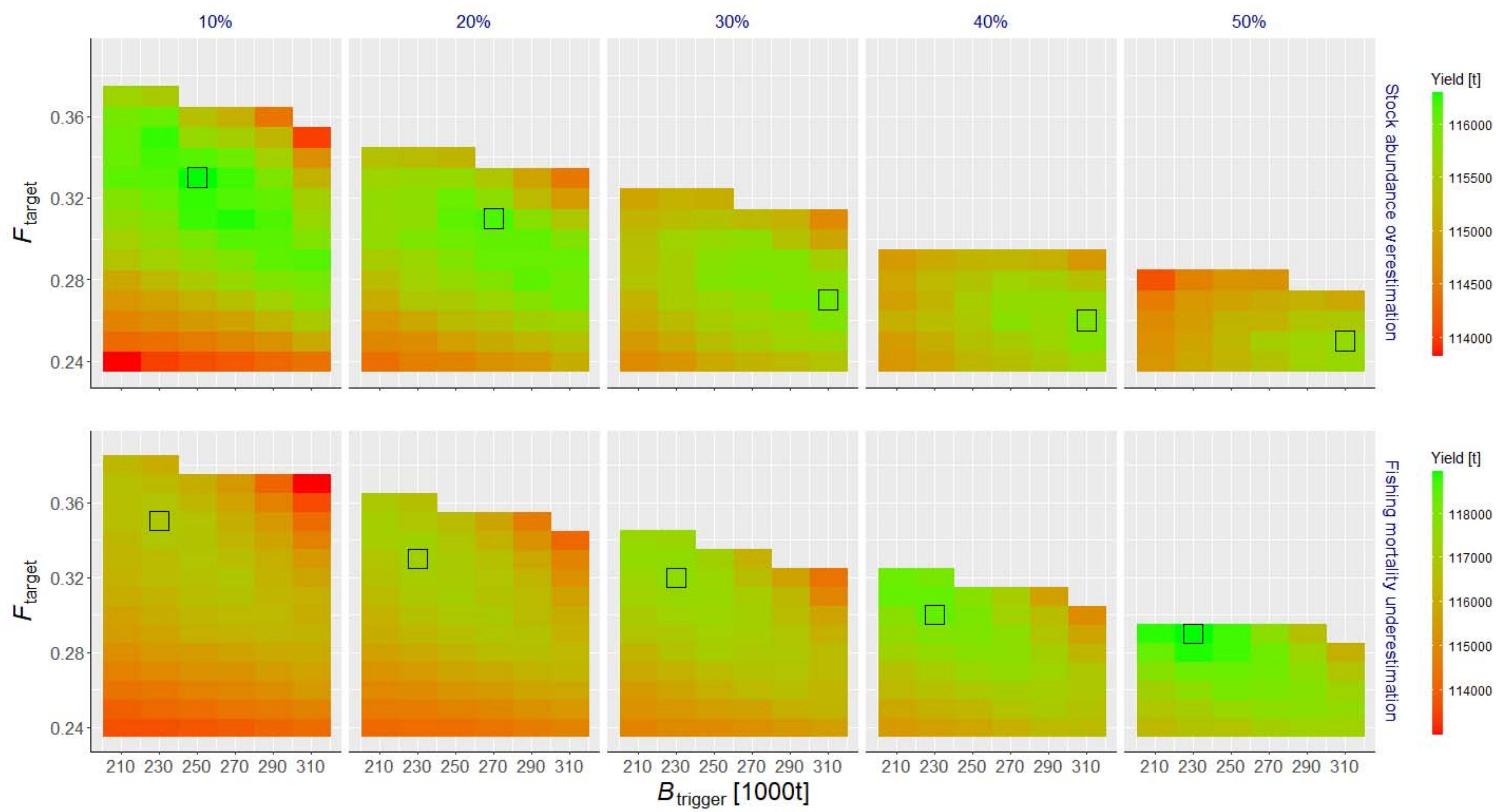

(a)
(b)
(c)
Catch [1000 t]
SSB [1000 t]
Interannual catch variability
Risk
base
+10N
+20N
+30N
+40N
+50N
-10F
-20F
-30F
-40F
-50F

Figure 4.


10%
20%
30%
40%
50%
Stock abundance overestimation
Fishing mortality underestimation
Yield [t]
116000
115500
115000
114500
114000
Yield [t]
118000
117000
116000
115000
114000
F_{target}
0.36
0.32
0.28
0.24
210 230 250 270 290 310
$B_{trigger}$ [1000t]