Received: 14 March 2019 | Revised: 7 May 2019 | Accepted: 14 May 2019

DOI: 10.1111/faf.12382

ORIGINAL ARTICLE

FISH and FISHERIES WILEY

Misspecification in stock assessments: Common uncertainties and asymmetric risks

Adrian R. Hordyk | Quang C. Huynh | Thomas R. Carruthers

Institute for the Oceans and Fisheries, University of British Columbia, Vancouver, British Columbia, Canada

Correspondence
Adrian R. Hordyk, Institute for the Oceans and Fisheries, University of British Columbia, 2202 Main Mall, Vancouver, BC V6T 1Z4, Canada.
Email: a.hordyk@oceans.ubc.ca

Funding information
Fisheries and Oceans Canada; Natural Resources Defense Council

Abstract

Common uncertainties in stock assessment relate to parameters or assumptions that strongly determine both the estimates of quantities of management interest (e.g. stock depletion) and related reference points (e.g. biomass at maximum sustainable yield). The risks associated with these uncertainties are often presented to managers in the form of decision tables. However, a formal evaluation of the risks from misspecifying an assessment model over time-horizons spanning multiple assessment cycles requires closed-loop simulation. There were two aims of this study: (a) develop an approach to identify and evaluate asymmetries in risk to yields and spawning biomass due to biases in key parameters and data sources in a stock assessment model, (b) quantify the relative importance of correctly specifying the various assessment attributes. A computationally efficient stock reduction analysis was evaluated using closed-loop simulation to identify risks associated with a stock assessment with persistent positive and negative biases in the key parameters and inaccurate assumptions regarding data sources. Six types of assessment misspecification were examined, namely the assumed natural mortality rate, the assumed recruitment compensation ratio, the assumed age of maturity, a hyper-stable or hyper-deplete index of abundance, over- or under-reporting of historical catch, and misspecification of the assumed shape of the selectivity curve. This study reveals large asymmetries in risk associated with common uncertainties in stock assessment processes. We highlight the value of reproducible and computationally efficient stock assessment models that can be investigated by closed-loop simulation before being used for fisheries management.

KEYWORDS

catch reporting, closed-loop, hyper-stability, natural mortality, simulation modelling, steepness

1 | INTRODUCTION

Quantitative stock assessments typically estimate biological reference points and the current stock status relative to these reference points to inform management decision-making. Stock assessment involves fitting population dynamics models to data, such as annual catch and effort, indices of population abundance and catch composition by length class or age (Hilborn, 2003). Although there have been significant developments in the science of stock assessment, considerable assessment uncertainties can still persist in even the most well-studied fisheries (Maunder & Piner, 2015; Maunder & Punt, 2013; Stewart & Martell, 2014).

Common stock assessment uncertainties relate to parameters or assumptions that can strongly determine both the estimates of quantities of management interest (e.g. stock depletion) and related reference points (e.g. biomass at maximum sustainable yield [MSY]). For example, the degree of recruitment compensation (typically parameterized as steepness *h*; the fraction of unfished recruitment obtained when spawning biomass has been reduced to 20% of the unfished level) is a principal determinant of the resilience of a population to fishing and defines reference points such as the level of stock depletion where yields are maximized (Mangel et al., 2013). However, it is generally the case that recruitment compensation is poorly estimated by stock assessment models, and therefore, either an informative prior is prescribed for the parameter controlling recruitment compensation or it is fixed at the best estimate obtained by meta-analysis or other comparable studies (Forrest, McAllister, Dorn, Martell, & Stanley, 2010; Lee, Maunder, Piner, & Methot, 2012; Zhu et al., 2012). Similarly, the natural mortality rate (*M*) determines population productivity and resilience to fishing but is also poorly informed by typical fishery data (Clark, 1999; Kenchington, 2014), and is therefore often informed by meta-analytical methods of life history correlates (Hamel, 2015).

Uncertainty also exists in the data sources that are used in the stock assessment. For example, an index of abundance, generated from a scientific survey or fishery-dependent catch per unit effort (CPUE) data, is often used to inform changes in population abundance over time. While indices of abundance are typically assumed to be proportional to changes in the biomass, they may in fact mask (hyper-stability) or over-exaggerate (hyper-depletion) the real changes in stock biomass (Hilborn & Walters, 1992). Similarly, while catch-at-age or -length data are assumed to be representative of the age- and size-structure of the population, it can be the case that older and larger individuals are under-represented in the sample due to a dome-shaped selectivity curve (Punt, Hurtado-Ferro, & Whitten, 2014).

Meta-analyses and other comparative studies have been used to provide scientific rationale for decisions in stock assessment (Maunder & Punt, 2013; Punt & Dorn, 2014). However, stock assessment scientists are often presented with several equally plausible alternative hypotheses that may require discrimination using semi-quantitative or subjective reasoning. The choice of value of fixed parameters and interpretation of data sources will influence the results of a stock assessment, including both the value and precision of quantities of management interest such as current stock size and reference points (Hilborn & Walters, 1992). Often these reference points relate to MSY to frame stock status (e.g. current spawning biomass relative to the equilibrium spawning biomass that would produce MSY; B_{MSY}) and the fishing mortality corresponding with MSY (F_{MSY}).

Decision tables provide a way to present the outcomes of management options across a range of uncertainties or "states of nature" that can include alternative values of fixed parameters (e.g. *M* and *h*) or alternative interpretations of data (Punt & Hilborn, 1997). For example, the most recent stock assessment for U.S. Pacific arrowtooth flounder (Sampson et al., 2017) provided a decision table with alternative total allowable catches (TACs) by row

and three levels of natural mortality rate as states of nature by column. The table cells are populated with the resulting spawning biomass and stock depletion arising from each management option over a short time horizon. Such decision tables highlight the trade-offs that are common in fisheries management, for example, an increase in short-term yield may be accompanied by a higher risk of the stock declining to an undesirably low level. Based on such tables, decision makers may weigh up the expected consequences of the different assumptions and choose a management recommendation (e.g. a TAC) intended to guide the fishery towards achieving management goals.

Decision tables reveal the risks across the states of nature by comparing their respective outcomes. For example, the U.S. Pacific arrowtooth flounder decision table shows a high risk of stock collapse if catches were taken according to a stock assessment assuming a high natural mortality rate while in fact the low natural mortality rate scenario was the true state of nature (Sampson et al., 2017). Alternatively, if catches were determined from an assessment assuming a low natural mortality rate while the true natural mortality rate was actually higher, the risk of stock depleting to low levels is reduced albeit with an increased risk of forgone yield due to underfishing (Sampson et al., 2017).

Decision tables generated from stock assessments are usually based on a simple projections of predetermined catch levels over a relatively short time horizon (typically less than 10 years). The relative risks of misspecifying the assessment among the various types of assumptions (e.g. M, h, selectivity) are often poorly quantified, particularly over the longer term (Legault & Palmer, 2016). Quantifying these longer-term risks would be informative in fishery management settings where estimates of assessment parameters are uncertain and assumed values have changed over time. For example, the assumed value of h for U.S. South Atlantic red snapper changed from 0.75 to 0.85, then to 0.99 in the 2008, 2010 and 2017 stock assessments, respectively (SEDAR, 2008, 2010, 2017). Similarly, in the Gulf of Mexico, the assumed natural mortality rate at-age ogive for red snapper assessments was updated between 2005 and 2015 (SEDAR, 2005, 2015), where M at age 1 in the most recent year was 1.2 year^{-1}, up from 0.6 year^{-1}, while M at age 15 in 2015 was 0.081 year^{-1}, down from 0.1 year^{-1}. Let us assume that in these cases, the evidence to support these changes in assessment specification is not conclusive. It is possible that the potential benefits arising from updating the parameters are small relative to the costs if these changes are incorrect. However, as demonstrated in the decision table cited earlier, wrongly assuming a higher value of M when in fact the true value is lower has a higher risk of stock depletion than the reverse scenario (assuming M is lower than the true value). This is an example of risk asymmetry.

A risk asymmetry exists where an assumption about a parameter value in one direction results in higher risk, in this case risk of decreased biomass, compared to an equivalent assumption in the other direction. The asymmetry in the risk of depleted biomass and forgone yield associated with bias in stock assessment parameters such as M is widely recognized in fisheries science. For example, a yield-per-recruit analysis can be used to calculate F_{MSY} and B_{MSY}/B_0 with alternative M values and demonstrate that lower M values correspond with more biologically precautionary estimates of F_{MSY}. Similarly, the yield-per-recruit model can reveal the expected relative yield with different assumed values of M and often indicate that relatively high catches can still be achieved with chronic underfishing (i.e. "pretty good yield" Hilborn, 2010). However, yield-per-recruit analyses describe these risk asymmetries under equilibrium conditions and do not reveal how these risks change over time, nor how a persistent bias in assumed parameters impacts the interaction between the stock assessment and the stock dynamics over time.

A formal evaluation of the risks from misspecifying an assessment model over time-horizons spanning multiple assessment cycles requires closed-loop simulation, whereby an operating model simulates plausible scenarios for stock and fishing dynamics from which fishery data are generated. These simulated data are passed to the assessment model to provide management recommendations, (e.g. catch advice) that impact the simulated dynamics of the operating model in the future time-steps. These steps are repeated for many time-steps to quantify the theoretical performance of the assessment accounting for feedbacks between model fitting and the system dynamics. Accounting for this feedback and the estimation of the stock assessment is critical because many stock assessments may prove adaptable and provide appropriate advice even when assuming biased parameter values.

A simple and well-documented example is the classic Schaefer surplus production model by which carrying capacity K and intrinsic rate of increase r may be estimated. Fishery data are often well explained by a negatively correlated set of these parameters, trading carrying capacity K for productivity r but leaving estimates of maximum sustainable yield (MSY = $rK/4$), and therefore the recommended catch advice, relatively constant and precise (Hilborn & Walters, 1992). Closed-loop simulation is central to the management strategy evaluation (MSE) approach increasingly used by resource managers to evaluate alternative management policies (Butterworth & Punt, 1999; Cochrane, Butterworth, De Oliveira, & Roel, 1998; Punt, Butterworth, de Moor, De Oliveira, & Haddon, 2016). Since the true state of nature represented by the operating model is known exactly, it is possible to quantify both the short-term and long-term risks of misspecifying an assessment model.

The aims of this study were to (a) develop an approach to identify and evaluate asymmetries in risk to yields and spawning biomass due to biases in key parameters and uncertainties in data sources in a stock assessment model and (b) quantify the relative importance of correctly specifying the various assessment attributes. The goal was to inform the fishery assessment process when there is little empirical evidence to distinguish between alternative assessment assumptions that carry strong asymmetry in risk. To evaluate asymmetry in risk, we calculate how the catch and depletion generally change as the magnitude of bias in the parameters and data provided to a stock assessment model increase. If catch and depletion are sensitive to a biased parameter or data source, then that parameter or data type has risk associated with it. An asymmetry in risk occurs when the risks due to a positive bias and negative bias of the same magnitude are not equal.

2 | METHODS

2.1 | Simulation framework

Simulation of the fishery dynamics was carried out using the "OMx" standard of age-structured operating models included in DLMtool (Carruthers & Hordyk, 2018a, 2019) an open-source package developed within the R environment (R Core Team, 2018) for efficient closed-loop testing of fishery management procedures (e.g. a stock assessment). The code and data that support this study are openly available in the lead author's GitHub repository at https://github.com/AdrianHordyk/Risk_Asymmetry.

2.2 | Case-study operating models

Operating models were developed based on four recently assessed stocks: U.S. arrowtooth flounder (*Atheresthes stomas*, Pleuronectidae) (Sampson et al., 2017), Pacific hake (*Merluccius productus*, Merlucciidae) (Berger, Grandin, Taylor, Edwards, & Cox, 2017), U.S. Pacific ocean perch (*Sebastes alutus*, Sebastidae) (Wetzel, Cronin-Fine, & Johnson, 2017) and Silver warehou (*Seriolella punctata*, Centrolophidae) (Tuck, 2016). These stocks were chosen to represent a wide range of life

TABLE 1 The life history for the four case-study stocks used in the simulation analysis

Common name	Species	Natural mortality (M, year^{-1})	von Bertalanffy growth parameters		Recruitment	
			L_∞ (cm)	K (year^{-1})	Steepness (h)	Recruitment error (σ_R)
Arrowtooth flounder	*Atheresthes stomas*	0.26	70	0.17	0.9	0.72
Pacific hake	*Merluccius productus*	0.22	50	0.36	0.86	1.52
Pacific ocean perch	*Sebastes alutus*	0.05	42	0.169	0.5	0.7
Silver warehou	*Seriolella punctata*	0.3	50	0.31	0.75	0.55

Note: Recruitment is defined with a Beverton–Holt stock–recruit relationship using the steepness parameter (h), with a log-normally distributed process error with standard deviation σ_R.

history parameters, particularly the natural mortality rate and steepness parameters. Life history and fishing parameters were based on the maximum likelihood estimates from the stock assessments, with some modifications to provide greater generality in the interpretation of results (Table 1). For example, growth, natural mortality and selectivity were assumed to be stationary over time, and depletion at the end of the historical simulations was set to $0.5B_{MSY}$ for each stock. Where values were estimated for both sexes, the female parameters were used. Where selectivity was reported in time blocks, the most recent selectivity curve from the most significant fishing fleet was used. The selectivity pattern was asymptotic in all four operating models unless otherwise specified (Figure 1).

2.3 | Simulated conditions

The historical conditions of the stocks were simulated from unfished conditions to the specified depletion level ($0.5B_{MSY}$) over a 50-year time period. The stocks shared a common historical fishing pattern,

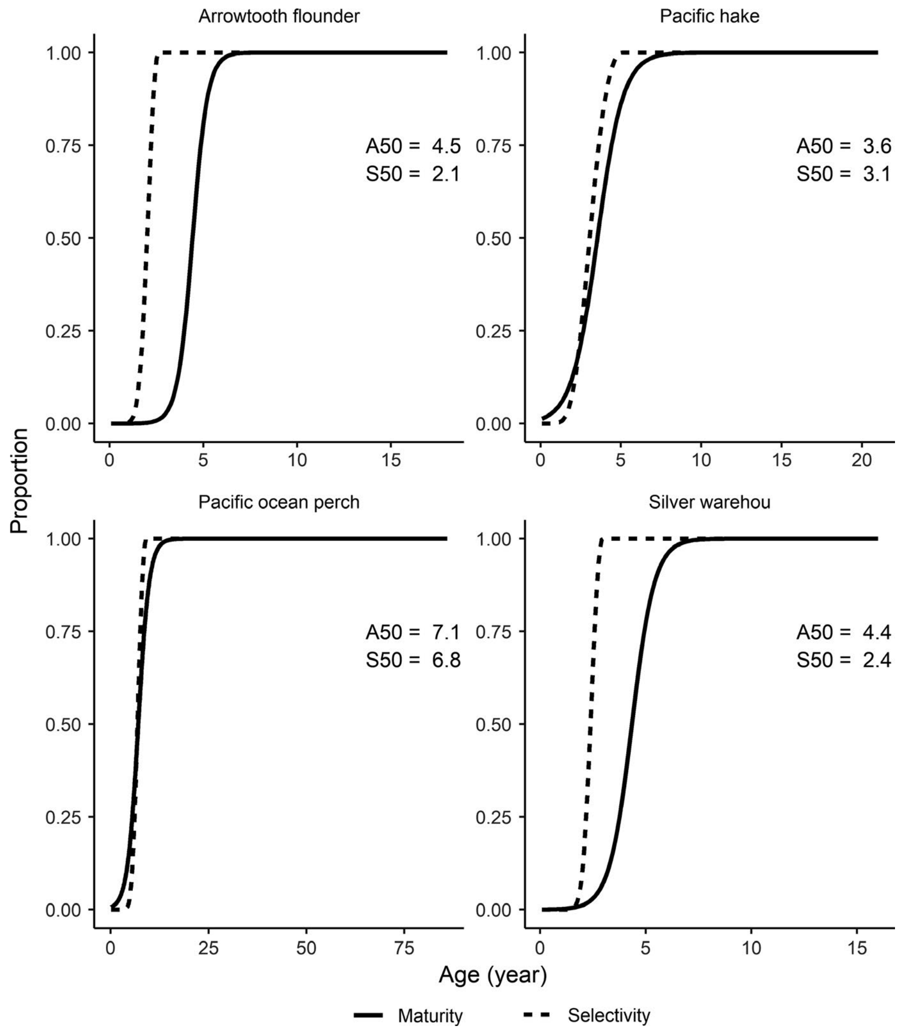

FIGURE 1 Maturity- (solid line) and selectivity-at-age (dashed line) for the four case-study stocks used in the analysis. The age (years) where 50% of the population is mature (A_{50}) and selected by the fishing gear (S_{50}) is shown in the top right corner for each stock

with fishing effort increasing for the three-quarters of the 50-year historical period and remaining constant for the last 12 years. For each simulation, the model optimized the catchability parameter q to drive the stock to the specified depletion level, given the historical fishing effort pattern, the random recruitment deviations, and the life history and selectivity parameters. Consequently, although primarily determined by the historical effort pattern, the simulated catch for the historical period was stock and simulation specific. The populations were then projected forward for 30 years, with an assessment model applied every 4 years to the simulated fishery data to obtain a TAC that was then applied to the stock in each of the next 4 years.

A total of 100 simulations were conducted for each operating model, type of assessment model assumption (e.g. bias in M; see description below) and value of that assumption. Age composition data was generated in each year (both historical and future projections) with an effective sample size in each simulation drawn from a uniform distribution spanning 50–100 individuals. Except where otherwise specified, the index of abundance was assumed to be proportional to vulnerable abundance, with log-normally distributed observation error with coefficient of variation (CV) for each simulation drawn from a uniform distribution ranging from 0.1 to 0.25. The same historical and future recruitment deviations and randomly drawn observation error parameters were used for evaluating each type of assessment model assumption. No implementation error was assumed for the TAC.

2.4 | Stock assessment model

An age-structured stock reduction analysis (SRA; Walters, Martell, & Korman, 2006) model was developed with Template Model Builder (TMB; Kristensen, Nielson, Berg, Skaug, & Bell, 2016) and used to perform the stock assessment (Huynh, Hordyk, & Carruthers, 2018). This model is fitted to an index of biomass and catch-at-age composition data (for details on how these data are simulated in closed-loop testing see Carruthers & Hordyk, 2018a) and estimates time-invariant selectivity and process error in the form of recruitment deviations. The TAC was set by multiplying the estimated exploitation rate corresponding to MSY by the current estimate of vulnerable biomass (*VB*). The parameters for M and h were fixed in the assessment model, and the selectivity pattern and unfished recruitment, the parameter that determines absolute stock size, were estimated. The stock assessment model is described in more detail in the Appendix S1.

2.5 | Description of analysis

The analysis quantified short-term and long-term risks to yield and biomass for six classes of uncertainties in the assumed parameters and data of the stock assessment model, specifically: a persistent bias in (a) the assumed values for natural mortality rate, (b) the assumed values for the steepness parameter, (c) the assumed values for the age of maturity; (d) the relationship between the index and the true stock biomass (i.e. hyper-stability and hyper-depletion); (e) over- or under-reporting of the catch, and (f) the shape of the descending limb of the selectivity curve in the assessment model (i.e. logistic to varying degrees of dome-shaped selectivity). Risk was defined as the change in yield and biomass compared to a simulation where the assessment was provided with unbiased parameters and representative data. To evaluate whether risks were asymmetric with respect to misspecification, for each class of uncertainty a range of values was assumed that spanned negative and positive biases.

Since M is a rate parameter, bound to positive values, and typically assumed to be log-normally distributed, bias was calculated in terms of the natural logarithm of M (log(M)). Two common methods to estimate M are the regression methods of Pauly (1980) and Hoenig (1983) (see also Then, Hoenig, Hall, & Hewitt, 2015, for updated versions of the Hoenig method). MacCall (2009) reports a standard error for log(M) of 0.53 and 0.5 for the Pauly and Hoenig methods, respectively. Based on these values, a maximum bias of ±50% was selected for log(M), approximating the 20th and 80th percentiles of a log-normally distributed variable with a standard deviation 0.5. Nine values of bias in log(M) were used in the analysis, ranging from −0.5 to 0.5 in increments of 0.125; that is, the analysis was run nine times with different assumed values of M provided to the stock assessment model, ranging from $Me^{-0.5}$ to $Me^{0.5}$, with an unbiased estimate of M in the fifth run of the model. For consistency, the same nine levels of bias were used for the other analyses.

To simulate bias in assumed recruitment compensation, the steepness parameter h, which assuming a Beverton–Holt stock–recruit model is bound between 0.2 and 1, was first converted to Goodyear's (1977) compensation ratio (CR) by $\mathrm{CR} = \frac{4h}{1-h}$, the bias applied to the CR, and then converted back to h which was provided to the stock assessment model. For example, a steepness value of 0.7 is equal to CR = 9.33, which with a −0.5 and 0.5 bias results in assumed CR of 4.66 and 14.0, or assumed steepness values of 0.54 and 0.78 respectively.

The maturity-at-age ogive is used to map the numbers-at-age to spawners-at-age in an age-structured populations dynamics model and therefore is important for determining the impact of fishing on the spawning biomass as well as estimated reference points such as F_{MSY}. Maturity ogives are typically estimated from maturity-at-age or -length data and provided as an input to a stock assessment model. While it is relatively straightforward to estimate the parameters of a logistic curve from maturity data, differentiating between physical development of gonads and functional maturity, and the timing and frequency of spawning mean that considerable uncertainty often exists in the relationship between age or length and maturity (Lowerre-Barbieri, Brown-Peterson, et al., 2011; Lowerre-Barbieri, Ganias, Saborido-Rey, Murua, & Hunter, 2011). For example, the stock assessment of the West Atlantic Bluefin tuna (*Thunnus thynnus*, Scombridae) considers two states of nature: a low age of maturity (100% at age 5) and a high age of maturity (100% at age 13) (ICCAT, 2017).

The evaluation of the impact of misspecification of the maturity schedule was conducted following the approach described above for the M and h parameters, with nine values of bias ranging from

−0.5 to 0.5 applied to the log of age at 50% maturity (A_{50}). To illustrate, a 50% negative and positive bias for A_{50} = 3 results in assumed an age of maturity of 1.8 and 4.9 years, respectively. The slope of the maturity ogive remained constant in the assessment model.

Indices of abundance in assessment models are treated as unbiased indicators of the relative trends in population abundance. To evaluate violations of this assumption, the operating models were conditioned to generate an index of abundance ranging from "hyper-stable," where the index declines slower than the stock, to "hyper-deplete," where the index declines faster than the stock. Harley, Myers, and Dunn (2001) reviewed 209 data sets with CPUE and an independent index of abundance and found that hyper-stability was most common, with the shape parameter β typically between 0.6 and 0.9. Although less common, hyper-depletion (β > 1) was observed in some stocks, with some β values estimated at 2 or higher. Based on this information, the index of abundance (*I*) was simulated as proportional to the vulnerable biomass (*VB*) raised to the power of the bias parameter β. Bias in β was specified in log-space using the same nine values described above, which results in β values ranging from 0.61 to 1.65.

Although changes in catch reporting are likely to be gradual rather than an abrupt change, for simplicity, we assumed that reported catches were subject to under-reporting (negative bias) or over-reporting (positive bias) for the first two-thirds of the historical time period (i.e. first 34 years of the 50-year period), but more recent catches assumed to be accurately reported. Following the previous analyses, bias ranged from a 50% under-reporting to a 50% over-reporting of the historical catch.

The final analyses examined misspecification of the assumed shape of the selectivity curve, focusing on the extent to which older, larger individuals are selected by fishing. In order to evaluate effect of both negative and positive bias in selectivity of larger fish, the selectivity curves in the operating model were changed for this scenario from a logistic curve to a double-normal selectivity pattern with the vulnerability at maximum age (V_{max}) fixed to 0.65 (Figure 2). The assessment model was then run with the nine bias values applied to the assumed V_{max} (Figure 2). This scenario represents a case where the true selectivity curve is dome-shaped, that is, selectivity-at-age declines after the age of full selection, and the stock assessment scientists chose to fix the shape of the descending limb of the selectivity curve rather than estimate it within the stock assessment. With a positively biased V_{max}, the assessment assumes that the selectivity curve is less dome-shaped than the underlying operating model, while with a negatively biased V_{max} the stock assessment assumes that the older individuals are less vulnerable to fishing (Figure 2).

In all cases, we assumed that the biases and uncertainty in data sources persisted throughout the entire projection period. That is, we did not consider cases where over time the biased parameters were detected and re-estimated.

2.6 | Quantifying risk and asymmetry in risk

The biological risk was summarized in terms of spawning biomass relative to expected unfished spawning biomass (i.e. stock depletion). Catches were standardized to a reference yield in each simulation, calculated as the highest obtainable yield at the end of the 30-year projection under a fixed *F* policy. Risk of decreased yields and low biomass was calculated in each year for each level of bias by dividing the catch and spawning biomass in each simulation by the unbiased run of the model. Asymmetry in risk due to over-estimation was calculated by dividing the risk from the model run with the positively biased parameters by the risk from the run with the equivalent negative bias (e.g. risk from +0.5 bias divided by risk from −0.5 bias). This is demonstrated in more detail in the first part of the Results section.

3 | RESULTS

3.1 | Description of asymmetry in risk

The analysis of bias in the assumed natural mortality rate for the Pacific hake case-study is used to demonstrate the evaluation

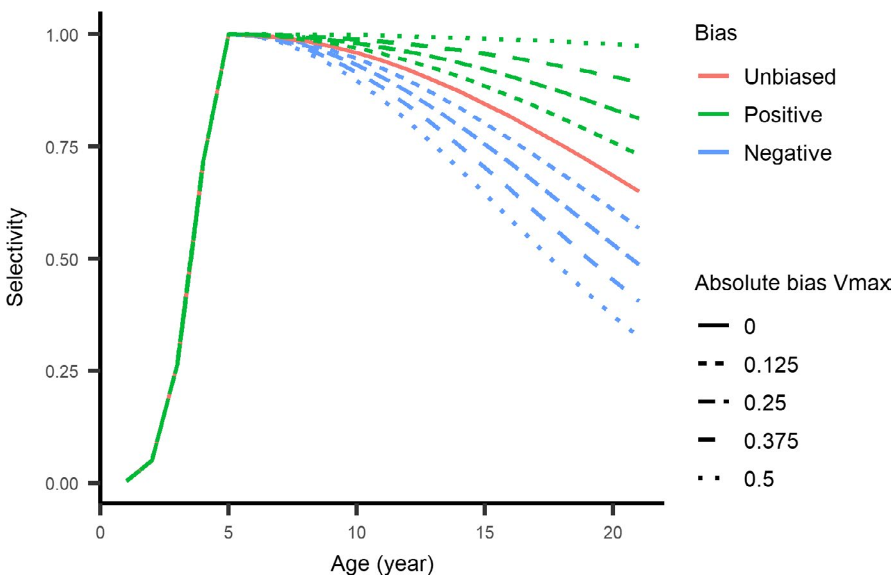

FIGURE 2 The dome-shaped selectivity curves used for the Pacific hake case-study for the analysis to evaluate impact of misspecification of selectivity. The nine levels of bias in the assumed vulnerability at maximum age (V_{max}) range from −0.5, where the assessment assumes selectivity is more dome-shaped than the operating model, to 0.5, where the assessment assumes selectivity is almost asymptotic, and includes bias = 0 where the selectivity pattern of the assessment matches the operating model (V_{max} = 0.65). Figure appears in colour in the online version only [Colour figure can be viewed at wileyonlinelibrary.com]

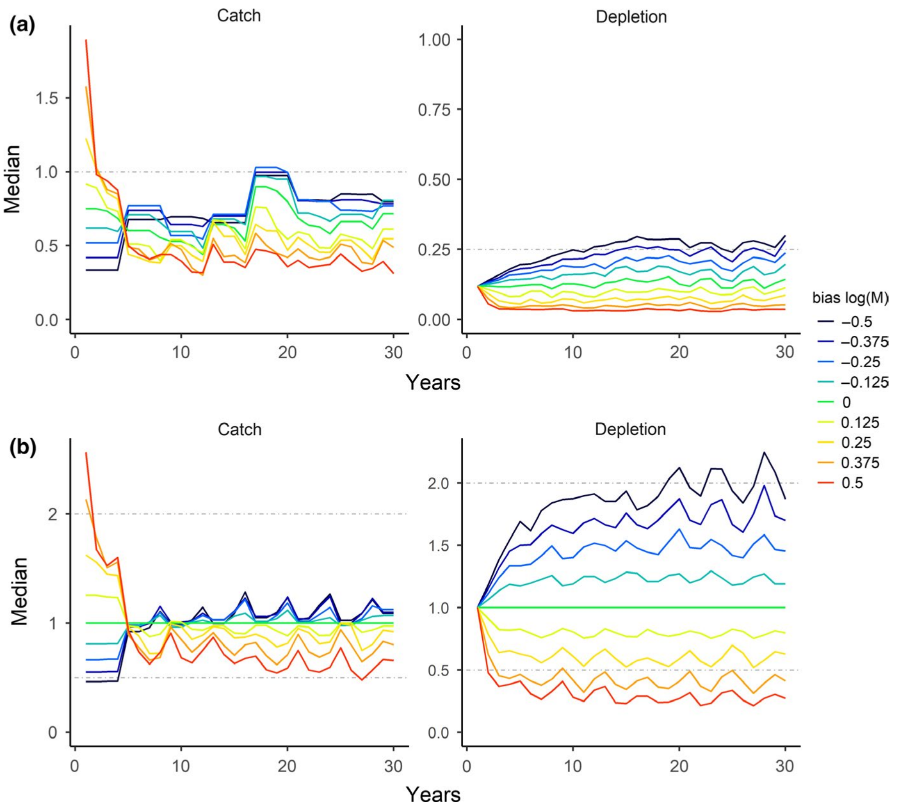

FIGURE 3 Projection plots for the Pacific hake case-study. In panel a, each line represents the median catch and depletion for each level of bias in log(*M*). The grey dash-dot lines indicate the reference yield and the depletion at B_{MSY}, respectively. In panel b, lines represent the median change in catch and depletion relative to the unbiased run of the model. Grey dash-dot lines mark the halving and doubling of catch and depletion relative to unbiased levels. Figure appears in colour in the online version only [Colour figure can be viewed at wileyonlinelibrary.com]

of asymmetry in risk of decreased yields and spawning biomass. Figure 3a shows the median catch and depletion for the nine levels of bias in log(*M*) over the 30 projection years. In the first year of the projection period, catches from runs with positively biased *M* are higher than the catches from the unbiased and negatively biased simulations, indicating that the short-term TAC recommendations were higher and lower when the assumed *M* was higher and lower than the true *M* in the operating model, respectively. Over time, catches were not sustained with positively biased *M* and subsequently decreased below those from the unbiased and negatively biased runs. For example, the median catch in the initial year was highest with the largest over-estimate of *M* (0.5 bias in log(*M*), Figure 3a) but generally declined throughout the projection period, resulting in the lowest median catch in the final year. In contrast, the simulations with the largest under-estimate of *M* (−0.5 bias in log(*M*); Figure 3a) had the lowest catch in the first four years. Median catch generally continued to increase after the second assessment (year 4) and in the final year was at a similar level to the −0.375 and −0.25 bias runs (Figure 3a).

All simulations started at the same level ($0.5B_{MSY}$) but there were immediate differences between the runs with different bias with respect to the median depletion (Figure 3a). Median depletion values in the unbiased run of the model (green line) stayed relatively stable throughout the projection period, well below B_{MSY}/B_0 of 0.25 (Figure 3a). The slow rebuilding time under the model run with unbiased *M* is not unexpected, as fishing mortality must be considerably lower than F_{MSY} to achieve rebuilding for overfished stocks (Carruthers & Agnew, 2016). All simulations with negatively biased *M* resulted in the stock increasing to higher levels compared to the unbiased run. The reverse pattern occurred with positively biased *M*, with the median depletion declining very quickly to 4% of unfished biomass when *M* had the largest positive bias (Figure 3a).

The focus of this study is risk asymmetries, and therefore, the absolute catch and depletion level are of less interest than the loss/gain in catch and biomass as positive and negative bias in assumed *M* increases. Figure 3b shows the median catch and depletion for each level of bias relative to the simulations with the unbiased estimate of *M*. The median catch in the first year was over twice of that from the unbiased simulation when *M* was positively biased by the largest amounts (0.375 and 0.5 bias in log(*M*)) and approximately half of the unbiased median catch when *M* was negatively biased by the largest amounts (−0.375 and −0.5 bias in log(*M*); Figure 3b). This pattern was reversed in the final years of the projection period, where positively biased *M* resulted in median catches ranging from 0.66 to 0.97 of the median catch from the unbiased run and the assessments with negatively biased *M* resulted in median catches closer to the unbiased catch in the final projection year (range 1.07–1.13 the median catch from the unbiased run; Figure 3b).

The relative difference in median depletion was more pronounced than the catches. For example, when the assessment was provided the largest positive bias in *M* the median spawning biomass in year 30 was almost a quarter of the depletion level in the unbiased run (Figure 3b). Conversely, when *M* was negatively biased by the same degree, the median spawning biomass continued to increase throughout the projection period and in the final year was 1.8 times higher than the biomass in the unbiased run (Figure 3b).

The risk of decreased yields and spawning biomass depended on the magnitude of the bias in log(*M*) and varied throughout the projection period. Risk of lower yields was higher in the short term when *M* was negatively biased and in the long term when *M* was positively biased (Figure 3). There was a less marked difference in risk of decreased spawning biomass over the short term and long term, with positive bias in *M* consistently resulting in a lower risk of the spawning stock falling to low levels.

Figure 3 only shows the trends in median catch and depletion and masks the variation within simulations. Figure 4a,b shows boxplots of the relative change in catch and depletion, respectively, for all 100 simulations and each level of bias in log(*M*) in the short term (year 4) and long term (year 30). Variability in short-term catches increased as bias increased from negative to positive values (Figure 4a). The reverse was true for depletion, where variability was highest when *M* was negatively biased and relatively constant across various degrees of positive bias (Figure 4b).

To calculate the asymmetry in risk associated with error in *M*, the relative catch and depletion for each simulation with a positive bias was divided by the relative catch and depletion from the corresponding negatively biased simulations. For example, the ratio of catch and depletion from the simulations where bias in log(*M*) = +0.125 was calculated, on a simulation by simulation basis, by dividing by the catch and depletion from the simulations where bias in log(*M*) = −0.125. Since the same recruitment deviations and observation errors were used for each model run, this comparison results in a catch and depletion ratio showing the expected loss (or gain) in catch and spawning biomass from a positive bias in natural mortality relative to a negative bias of the same magnitude, which is displayed for the short term and long term (years 4 & 30) (Figure 4c,d for catch and depletion ratios, respectively). A symmetrical risk profile, where the risk of decreased catch and spawning biomass with a positively biased *M* is equivalent to the risk with a negatively biased *M* of the same magnitude, would result in a ratio of one. Ratio values

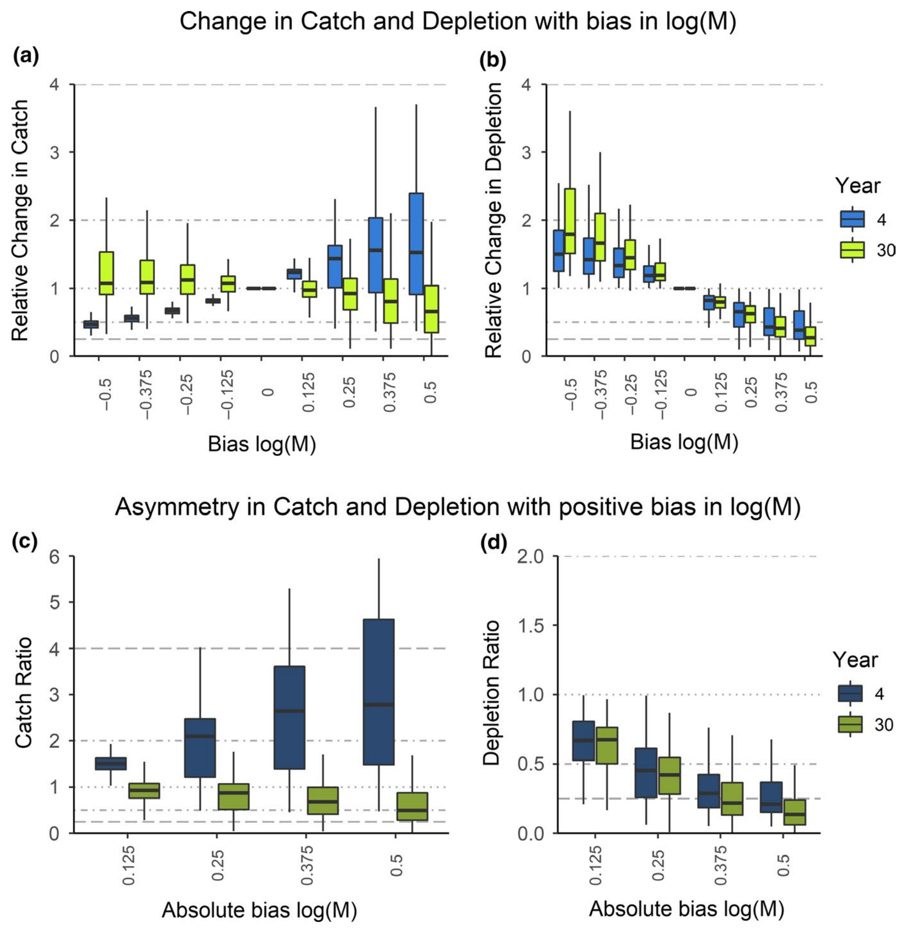

FIGURE 4 Boxplots showing the (a) relative change in short- and long-term catch and (b) depletion relative to an assessment with unbiased parameters for nine values of bias in log(*M*) for the Pacific hake case-study. Grey horizontal lines are added at one (dotted), and 0.5 and 2 (dashed-dot) and 0.25 and 4 (dashed) to indicate no relative change, half and double, and quarter and quadruple changes, respectively. Values greater than 1 indicate higher catch or stock biomass relative to the assessment with unbiased *M*. Panels (c) and (d) show the risk asymmetry in catch and depletion resulting from positively biased *M* and are calculated by dividing the catch and depletion from the simulations with a positive bias by the catch and depletion from the simulations with the equivalent negative bias. Values greater than 1 indicate higher catch or stock biomass in the positive bias assumption compared to the corresponding negative biased assumption. Boxes span the interquartile range and the whiskers span the 95% interval of observed values. Risk asymmetry is indicated by medians which deviate from 1 in panels (c) and (d). Figure appears in colour in the online version only [Colour figure can be viewed at wileyonlinelibrary.com]

greater than one indicate a risk asymmetry, with larger increases in catch/depletion with positive biased values than decreases with equivalent negative biased values. Similarly, ratio values less than one indicate the opposite risk asymmetry where positively biased *M* results in lower catch and depletion compared to an equivalent negative bias.

In the short term, positively biased *M* resulted in higher catches, indicating that negatively biased *M* results in a higher risk of decreased yields (Figure 4c). This pattern was reversed in the long term, where positively biased *M* was, on average, almost risk neutral with respect to decreased yields with a ±12.5% bias in log(*M*) and risks of lower yields increasing as the bias increased (Figure 4c). The pattern in risk asymmetry of decreased biomass remained the same in the short term and long term. In almost all cases, an assessment with a positive bias in log(*M*) resulted in lower levels of biomass in both the short term and long term compared to model runs with a negative bias of the same magnitude (Figure 4d).

The next section of the Results describes the relative impact of the six classes of uncertainties examined in this analysis in terms of the median catch and depletion ratios for each level of bias. Then, the following sections present a more detailed description of the analyses for each assessment uncertainty and case-study. The results focus on the median catch and depletion ratios for the 100 simulations. Boxplots showing the distribution of the relative change in catch and depletion, and the catch and depletions ratios (following the format shown in Figure 4) are shown for each case-study and type of assessment misspecification in the supplementary material (Figures S1–S12).

3.2 | The relative impact of types of assessment misspecification

Bias in natural mortality had the largest asymmetry in risk in terms of both short-term and long-term catches (Figure 5). Short-term median catches were typically 2–4 times higher when *M* was positively biased by the largest amount compared to the catches from a negatively biased *M* of the same magnitude. With the exception of the Pacific ocean perch case-study, long-term catches were lower when *M* was positively biased, with catches less than half and some cases much less than a quarter compared to the long-term catches from

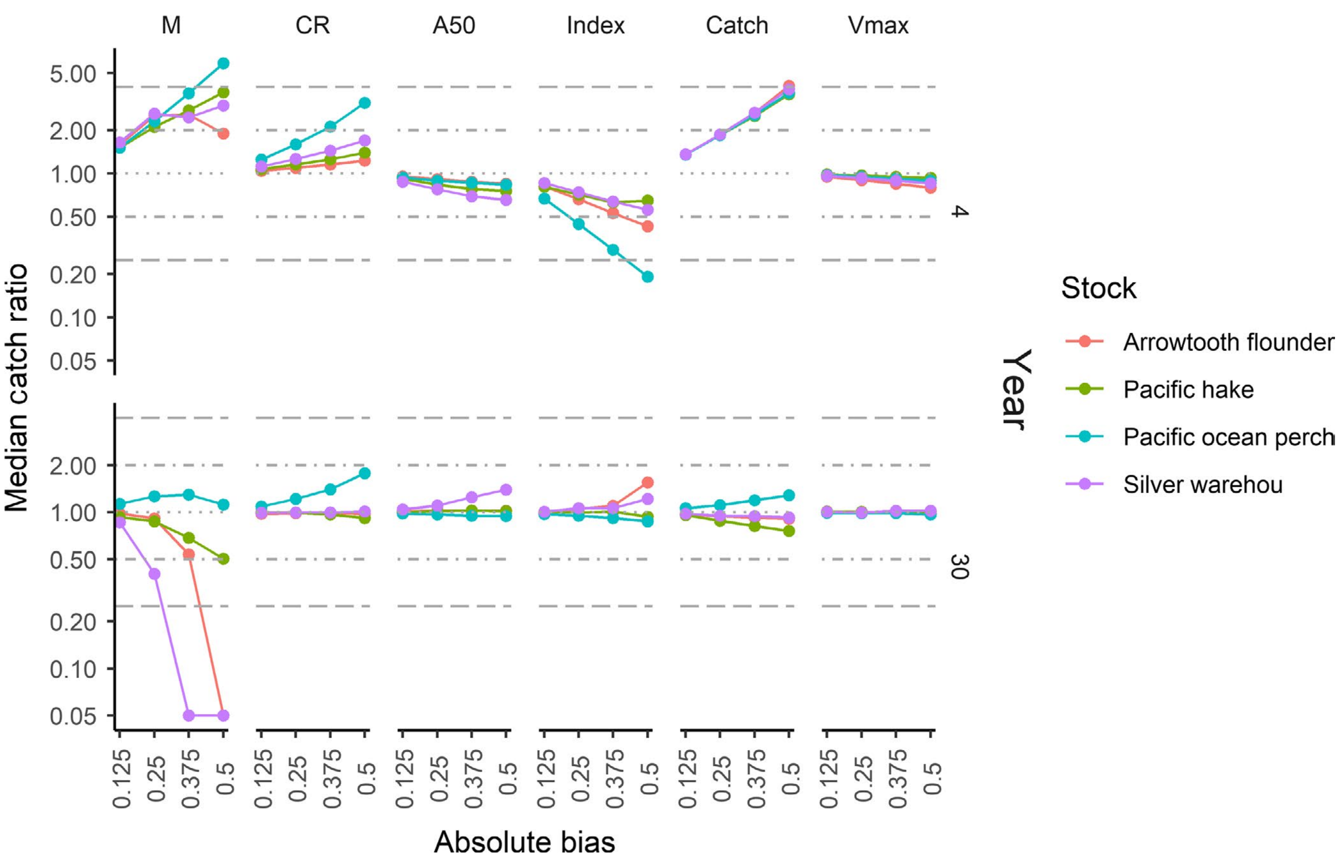

FIGURE 5 The median short- (top row) and long-term (bottom row) catch ratios for the four case studies and the six variables examined in this study (*M*: the assumed natural mortality rate, CR: the assumed recruitment compensation ratio, A_{50}: the assumed age of maturity, Index: a hyper-stable or hyper-deplete index of abundance, Catch: over- or under-reporting of historical catch, and V_{max}: the assumed vulnerability of the oldest age-class). The plot reveals the relative difference in the asymmetry of risk of decreased yields as a consequence of positive bias in these variables and highlights that the risk profile is species-specific in some cases. Grey horizontal lines are added at one (dotted), and 0.5 and 2 (dashed-dot) and 0.25 and 4 (dashed) to indicate no relative change, half and double, and quarter and quadruple changes, respectively. The scale for the *y*-axis is shown in log-space to aid with interpretation. Values have been truncated to a minimum value of 0.05 for display purposes. Figure appears in colour in the online version only [Colour figure can be viewed at wileyonlinelibrary.com]

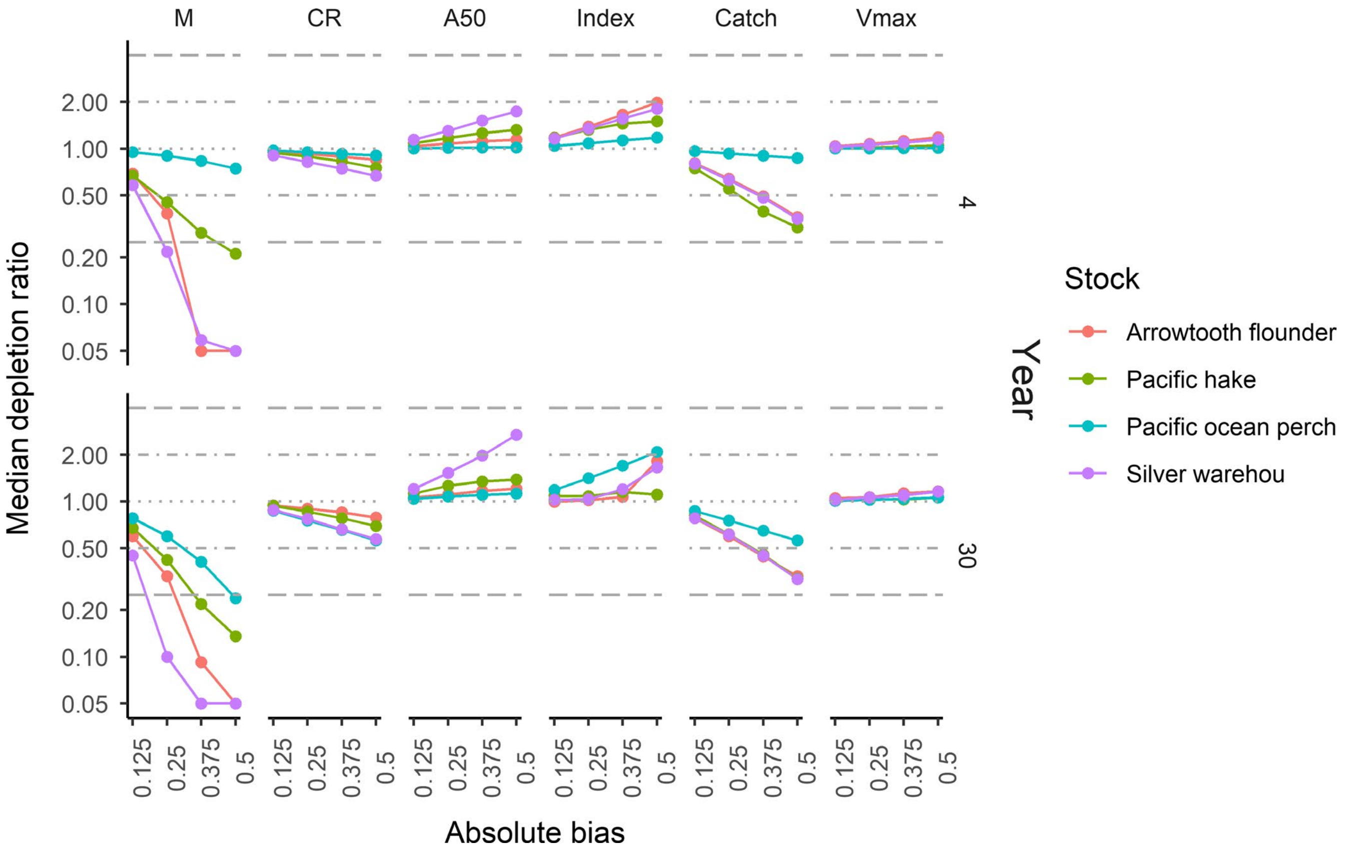

FIGURE 6 The median short- (top row) and long-term (bottom row) depletion ratios for the four case studies and the six variables examined in this study (*M*: the assumed natural mortality rate, CR: the assumed recruitment compensation ratio, A_{50}: the assumed age of maturity, Index: a hyper-stable or hyper-deplete index of abundance, Catch: over- or under-reporting of historical catch, and V_{max}: the assumed vulnerability of the oldest age-class). The plot reveals the relative difference in the asymmetry of risk of decreased spawning abundance because of positive bias in these variables and highlights that the risk profile is species-specific in some cases. Grey horizontal lines are added at one (dotted), and 0.5 and 2 (dashed-dot) and 0.25 and 4 (dashed) to indicate no relative change, half and double, and quarter and quadruple changes, respectively. The scale for the *y*-axis is shown in log-space to aid with interpretation. Values have been truncated to a minimum value of 0.05 for display purposes [Colour figure can be viewed at wileyonlinelibrary.com]

a negatively biased *M* value (Figure 5). A positive bias in the age-at-maturity resulted in lower short-term catches compared to the catch from negatively biased age-at-maturity; however, in the long term the catches were roughly symmetrical for all stocks except silver warehou, where long-term catches were higher when age of maturity was positively biased. Compared to other forms of assessment uncertainty, misspecification of selectivity (the vulnerability of older individuals) had symmetrical risks for the stock levels and catches over both the short term and long terms (Figure 5).

Bias in the assumed natural mortality rate and catch misreporting had the largest asymmetries in the risk to both short- and long-term biomass (Figure 6). At absolute bias values of 0.25 and greater, the median biomass from positively biased *M* was typically less than half, and in some cases, much less than quarter, of the stock biomass when *M* was negatively biased (Figure 6). To a lesser extent, a positive bias in the assumed recruitment compensation ratio also generated substantial risk to stock biomass. The direction of the risk asymmetry was reversed for the age of maturity (A50) and bias in the index of abundance (Index), where a positive bias resulted in biomass 1–2 times higher than the median biomass from negatively biased values (Figure 6). A positive bias in reporting of historical catches resulted in increased risk of stock depletion for all stocks, especially over the long term. The risk asymmetry in median depletion was least pronounced for the misspecification of selectivity.

3.3 | Natural mortality

In the short term (year 4 of the projection), the catch ratio typically was positive, indicating a risk asymmetry with positively biased *M* resulting in higher immediate catches compared to a negatively biases of same magnitude (Figure 5). Conversely, long-term catches (calculated in the final of the 30 projection years) decreased with positive bias in *M* for arrowtooth flounder, Pacific hake and silver warehou. Positive biases of 0.375 and 0.5 resulted in catches that were very low compared to the corresponding negative bias for both the arrowtooth flounder and the silver warehou (Figure 5). The exception was Pacific Ocean perch. For this stock, the median long-term catch ratio remained positive for all levels of bias (Figure 5). For this relatively long-lived stock, the 30-year projection was insufficient to evaluate long-term impacts of high yield on the age structure of the population.

The declines in projected biomass were more than compensated for by increases in prescribed fishing mortality rate, and the simulated populations were “mined” allowing yields to remain relatively high.

There was also asymmetrical risk with respect to stock depletion. Positive *M* bias resulted in considerably lower stock biomass for all four case studies (Figure 6). In the short term, a positive bias of 0.375 or greater resulted in spawning biomass declining to very low levels compared to that arising from a negative bias for the arrowtooth flounder and silver warehou stocks (Figure 6). The asymmetry in depletion was less pronounced for the Pacific hake, where a positive bias of 0.375 and 0.5 resulted in a short-term median depletion ratio of 0.29 and 0.22, respectively (Figure 6). The simulated Pacific ocean perch stock was least sensitive to positive *M* bias in the short term, with a median short-term depletion ratio of 0.95, 0.90, 0.84 and 0.75 for the four levels of bias (Figure 6). Arrowtooth flounder, the Pacific hake and the Silver warehou case studies showed similar patterns in long-term depletion risk with positive bias in *M*, with steep declines in relative biomass with increasing degree of bias (Figure 6). A risk asymmetry also existed in the long-term depletion ratio for the simulated Pacific ocean perch case-study, although this was not as pronounced as the other stocks (Figure 6).

3.4 | Recruitment compensation ratio

There was an asymmetric risk to catches over the short term when recruitment compensation ratio, CR, was misspecified for all four stocks, with higher catches in the short term when CR was positively biased compared to catches from a negatively biased estimate of CR (Figure 5). The asymmetry was most pronounced for the Pacific ocean perch, where the median catches from an assessment with positively biased CR were 1.25 to over three times higher than the catches from the assessment with an equivalent negatively biased steepness parameter (Figure 5).

The asymmetry in risk of bias in assumed CR was considerably lower on the long-term yield for the stocks (Figure 5). The median catch ratio was close to one for all levels of bias for all case studies except the Pacific ocean perch, where the catch ratio continued to increase as bias in recruitment compensation increased, indicating that higher long-term catches were achieved when CR was positively biased (Figure 5).

Both short- and long-term median spawning stock biomass were lower when CR was positively biased, with depletion generally lower in the long term (Figure 6). The Pacific ocean perch had the largest difference in short- and long-term depletion ratio, where the minimum depletion ratio with the most biased CR (bias = 0.5) was 0.90 and 0.56 in the short term and long term, respectively (Figure 6).

3.5 | Age of maturity

A positively biased assumed value for the age of maturity (A_{50}) resulted in short-term catches that were lower than those from a negatively bias, and the catch ratio tended to decrease as the level of bias increased (Figure 5). The asymmetry was most pronounced for the silver warehou, where a positive bias resulted in median catches that were 12.5%, 22%, 30% and 34% lower than those from an equivalent negatively biased estimate of A_{50} (Figure 5). There was an approximately symmetrical risk profile for long-term catch for all case-study stocks expect silver warehou, where increasingly positively biased A_{50} resulted in higher median long-term catches compared to the catches from a negatively biased A_{50} (Figure 5).

Negatively biased A_{50} generally resulted in increased spawning biomass in both the short term and long term (Figure 6). The Pacific ocean perch was an exception where increasing levels of bias in A_{50} had little impact on the depletion in the short term, and a marginal positive asymmetry in the long term (Figure 6).

3.6 | Index of abundance

A hyper-deplete index of abundance (positive bias in log(β)) resulted in lower short-term median catches for all stocks compared to a hyper-stable index with equivalent absolute bias (Figure 5). This was most pronounced for the Pacific ocean perch, where the short-term median catches from a positively biased index were 0.67, 0.45, 0.30 and 0.19 of the catches from a negatively biased index (Figure 5).

There was less of an asymmetry in risk for the long-term catches for the four cases studies, with a moderately positive asymmetry (higher catches) for the arrowtooth flounder and the silver warehou, and moderately negative asymmetry (lower catches) for the Pacific hake and Pacific ocean perch, especially at higher levels of bias (Figure 5). However, there was large variability in all stocks except for the Pacific ocean perch, with the tails of the distributions extending well beyond one and close to zero, indicating both higher and lower catches occur with a hyper-deplete index (see Figures S7 and S8).

There was a positive asymmetric risk profile for short-term depletion for all stocks, with a positive bias in the index of abundance resulting in higher short-term biomass compared to an equivalent negatively biased index (Figure 6). The Pacific ocean perch had the least asymmetry in risk of short-term depletion. In contrast, the hyper-stable index of abundance had the greatest impact on the long-term depletion of the Pacific ocean perch stock, with spawning stock biomass at the end of the projection period tending to increase as the bias in the hyper-deplete index increased compared to a hyper-stable index with the same absolute bias (Figure 6). Median long-term depletion for the other three stocks tended to stay relatively stable as the bias in index increased, except at the highest level of bias for arrowtooth flounder and silver warehou, where the median depletion ratio increased (Figure 6). However, similar to the catches described above, there was high variance in depletion outcomes, particularly for the arrowtooth flounder and silver warehou stocks, where for some simulations depletion fell to less than half, and in many cases less than a quarter, of the biomass from a hyper-stable index (see Figures S7 and S8).

3.7 | Reporting of catch

A positive bias in historical catches resulted in higher short-term catch for all stocks compared to a comparable under-reporting of catch (Figure 5). There was less of an asymmetry in risk to long-term catches, although long-term catch tended to be lower with a positive bias in historical catches compared to a negative bias of the same magnitude, except for the Pacific ocean perch, where long-term catches were higher with a positive bias in reported catch (Figure 5).

Positively biased historical catches always resulted in lower stock levels when compared to under-reporting of the same scale. In many cases, the stock in the over-reporting scenarios declined to a small fraction (e.g. less than half or close to a quarter) of the abundance in the corresponding under-reporting scenarios (Figure 6).

3.8 | Misspecification of selectivity

Misspecification of the shape of the selectivity curve resulted in asymmetric short-term yields, with a positive bias in the assumed vulnerability of the maximum age class (i.e. assuming the selectivity curve is less dome-shaped), resulting in lower short-term catches than an comparable negative bias for all four case studies (Figure 5). The risk profile for the long-term catches was more symmetric than the short-term (Figure 5). Similarly, both short- and long-term spawning abundance were typically higher when older individuals were assumed to be more vulnerable to the fishery (Figure 6).

4 | DISCUSSION

Establishing an objective, scientifically rigorous stock assessment is a stated goal of most fishery management frameworks. Even in cases where subjective judgement is relied on to conduct an assessment, care is generally taken to identify uncertainty in parameters and data sources. In most cases, these alternative states of nature are investigated through either sensitivity analysis or decision tables, which present the short-term trade-offs in catch and risk of stock decline for alternative assumptions in the model. While informative for short-term decision-making, the lack of feedback between the assessment model and the projected population means that decision tables are not suited to reveal the long-term risks of a persistent misspecification over the course of an assessment process. In contrast, the closed-loop simulation testing of the assessment model used in this study includes the feedback between the stock assessment model and the projected population, and therefore can account for learning that occurs over time.

The results of this study reveal large asymmetries in risk associated with common uncertainties in stock assessment processes. There was a much higher risk to long-term yields and stock biomass when specifying positively biased natural mortality rates, recruitment compensation and catch levels than the equivalent negative bias. Qualitatively, the direction of risk derived from closed-loop simulation in these analyses correspond with sensitivities reported by previous research investigating age-structure model estimates to varying natural mortality rates (Clark, 1999; Deroba & Schueller, 2013; Williams, 2002) and catch biases (Van Beveren et al., 2017; Omori, Hoenig, Luehring, & Baier-Lockhart, 2016; Rudd & Branch, 2017). Additionally, this analysis demonstrates that reconstructing historical catches to account for under-reporting (e.g. Pauly & Zeller, 2016) provides benefits in short-term yield but has the cost of increased risk to long-term biomass.

For natural mortality rates, recruitment compensation and catch levels, the asymmetry in risk to short-term biomass followed a similar pattern to long-term risk, increasing with positive bias (albeit lessened). However, a strongly contrary pattern emerged for short-term yields. For example, positive biases in natural mortality rate that would all but collapse a fishery in the long-run provided much higher yields over the short term. The Pacific ocean perch case-study confirms that for longer-lived species, it may take a much longer time horizon to realize the eventual costs of assessment misspecifications that favour short-term benefits. The trade-off between short and long-term yields is well established for exploited systems (Walters & Martell, 2004) and has been repeatedly demonstrated among management procedures (e.g. Carruthers et al., 2014). This is the first time that such stark trade-offs have been revealed for data-rich stock assessments across typical ranges of uncertainty for parameters and assumptions.

A principal concern arising from these results is that if core uncertainties prevail by objective methods, this could open the door for subjective influence. No evidence is presented in this paper to suggest subjectivity has influenced assessments in a deliberate attempt to achieve short-term yield at the cost of long-term risk to yields and biomass. However, the results demonstrate that prevailing assessment uncertainties span a wide trade-off between short-term and long-term outcomes of varying importance to typical stakeholder groups. Current approaches (e.g. decision analysis) provide a rigorous basis for investigating short-term benefits of assuming a particular state of nature, but because they do not account for the feedback between the misspecified assessment and the population dynamics, cannot adequately quantify the long-term costs of misspecification. This requires repeated implementation of the assessment in closed-loop simulation as demonstrated in this paper.

There may be sound scientific justifications for historical changes in parameters for certain stocks (for example the changes in recruitment compensation in U.S. South Atlantic red snapper previously cited), and the results of this study do not imply otherwise. However, the results indicate that if the relative role of subjectivity were considered high, it may be beneficial to review historical assumption-creep, particularly for natural mortality rate, recruitment compensation and catch reporting because they come with the highest asymmetry in risk to the sustainability of yields and biomass. Long-term biological and economic interests are significant management objectives in many countries with developed fishery legislation (e.g. the U.S. Magnuson-Stevens Fishery Conservation and Management Act). However, there may be legitimate grounds for the prioritization of short-term yields over long-term interests in certain management settings. It is straightforward to characterize those assessment processes most likely to favour short-term over long-term interests, for example those where empirical evidence is lacking, uncertainty is large, and restrictive management measures are contentious.

Knowledge of asymmetric risk may help managers direct oversight with respect to certain aspects of stock assessments. For example, where a proposed change to an assessment parameter implies a significant increase in risk, managers may place a greater burden of proof on analyses in support of the change. It may be advisable to conduct case-specific analysis of assessment misspecification rather than relying on more general studies such as this. Among the case studies of these analyses, there was a high degree of variability in the outcomes that would be increased substantially given alternative levels of stock depletion, assessment types, harvest control rules (we assumed F_{MSY} fishing levels rather than a more restrictive harvest control rule), data quality, duration of projections and fishery management objectives. While in this analysis misspecification of one assumption at a time was considered, a tailored analysis for a specific fishery might evaluate risks for simultaneous biases across multiple assessment assumptions (e.g. a poorly defined range of natural mortality rates and corresponding selectivities; Thompson, 1994).

Analyses of risk for a specific fishery could also examine several other uncertainties that we were not able to include in this study. For example, reliable age data are often lacking in some assessment processes, and in such cases, it would be valuable to evaluate the risk asymmetries associated with misspecification of the growth parameters. In this study, we examined the risk profile for single parameters and data sources. In most cases, uncertainties would exist in several parameters and data sources simultaneously, and the risk associated with the interaction between biases in several parameters could be examined with a study designed for this purpose. Further research could also investigate the impact of these biases and uncertainties in alternative assessment models. For example, while our assessment model used age composition data, further research could explore an assessment model that fits to length data.

A "best case" recommendation arising from this research is that stock assessments (and any harvest control rules that translate the assessment into management recommendations) should be reproducible and computationally efficient enough to be investigated by closed-loop simulation. A response to this proposition has been that stock assessments are inherently subjective and are intended to include a wide range of valuable expert judgement and therefore can never be adequately represented in simulation testing (Rice & Rochet, 2009). This raises practical and philosophical questions. On a practical level, if the assessment approach is not reproducible it will not be possible to evaluate expected performance and hence to justify one *modus operandi* over another (e.g. data collection, duration of the assessment cycle, assessment model, harvest control rule) (Butterworth et al., 2010). On a philosophical level, if the assessment process includes a high degree of subjectivity and it is accepted that another group of experts could derive a substantially different management recommendation, it may be hard to defend its scientific validity (e.g. "best scientific practice") for which objectivity and reproducibility are defining attributes.

Having identified the serious potential risks of assessment misspecification, a priority for future research is the development of suitable diagnostic tools. Current approaches include analysis of retrospective patterns and residual analysis (Carvalho, Punt, Chang, Maunder, & Piner, 2017; Hurtado-Ferro et al., 2015; Legault, 2009) and predictive skill (Brooks & Legault, 2015; Kell, Kimoto, & Kitakado, 2016). Currently, under investigation is the development of methods that compare the data observed after an assessment was conducted with the posterior predictive data of the assessment to detect persistent model misspecification (e.g. Carruthers & Hordyk, 2018b). Future research could include the post hoc analysis of assumption-creep in assessed fisheries to determine whether there have been risky changes in assumptions that are not well supported by empirical evidence.

The availability of efficient estimation routines (MSEtool: Huynh et al., 2018; TMB: Kristensen et al., 2016), closed-loop frameworks (DLMtool: Carruthers & Hordyk, 2018a), has unlocked the ability to easily and rapidly test the theoretical performance of proposed fishery management systems, including data-rich stock assessments, before they are implemented. In fisheries management, closed-loop simulation has traditionally focused on MSE to identify robust management procedures and reveal performance trade-offs (Butterworth & Punt, 1999; Punt et al., 2016). However, closed-loop simulation can inform a wider range of fishery management decision-making. In this research, it revealed the long-term outcomes of a consistently misspecified but potentially adaptable (via re-estimation) stock assessment process. Some traditional assessment uncertainties were demonstrated to be relatively unimportant (e.g. maturity, gear selectivity) compared to other uncertainties, potentially allowing stock assessment processes focus on the most consequential assessment uncertainties.

Closed-loop simulation testing has notable flaws. We have particular concerns over the simulation of data that are much better behaved than real fishery data (i.e. are simulated from similar assumptions to the observation models of assessments) and the consideration of a narrow range of structural uncertainties by using operating models based on conventional age-structured equations with scenarios based on "known unknowns" (a limitation of this analysis also). However, given that empirical management experiments are not feasible for most fisheries, we recommend (as have others; Butterworth et al., 2010; Cooke, 1999) that where possible, fishery management systems are formalized for testing by closed-loop simulation. Despite its limitations, it remains the only theoretically coherent framework available for justifying one proposed fishery management *modus operandi* over another and inform choices over data collection and processing, stock assessment model complexity, harvest control rules and the implementation of management measures.

ACKNOWLEDGEMENTS

We thank Carl Walters for helpful comments on an earlier draft of this manuscript. We also thank the editor Paul Hart, Andre Punt and an anonymous reviewer for helpful comments and suggestions that improved this work. AH acknowledges funding support from the Natural Resources Defense Council. QH was funded by Grant and Contribution

Funds (Ocean and Freshwater Science Contribution Program) from Fisheries and Oceans Canada for development of MSEtool.

DATA ACCESSIBILITY

The code and data that support this study are openly available in the lead author's GitHub repository at https://github.com/AdrianHordyk/Risk_Asymmetry.

ORCID

Adrian R. Hordyk https://orcid.org/0000-0001-5620-3446
Quang C. Huynh https://orcid.org/0000-0001-7835-4376
Thomas R. Carruthers https://orcid.org/0000-0001-9173-9043

REFERENCES

Berger, A., Grandin, C. J., Taylor, I. G., Edwards, A. M., & Cox, S. (2017). *Status of the Pacific Hake (Whiting) stock in U.S. and Canadian Waters in 2011*. Joint Technical Committee of the Pacific Hake/Whiting Agreement between the Governments of the United States and Canada.

Brooks, E. N., & Legault, C. M. (2015). Retrospective forecasting—Evaluating performance of stock projections for New England groundfish stocks. *Canadian Journal of Fisheries and Aquatic Sciences*, *73*(6), 935–950. https://doi.org/10.1139/cjfas-2015-0163

Butterworth, D. S., Bentley, N., De Oliveira, J. A. A., Donovan, G. P., Kell, L. T., Parma, A. M., … Stokes, T. K. (2010). Purported flaws in management strategy evaluation: Basic problems or misinterpretations? *ICES Journal of Marine Science*, *67*(3), 567–574. https://doi.org/10.1093/icesjms/fsq009

Butterworth, D. S., & Punt, A. E. (1999). Experiences in the evaluation and implementation of management procedures. *ICES Journal of Marine Science*, *56*(6), 985–998. https://doi.org/10.1006/jmsc.1999.0532

Carruthers, T. R., & Agnew, D. J. (2016). Using simulation to determine standard requirements for recovery rates of fish stocks. *Marine Policy*, *73*, 146–153. https://doi.org/10.1016/j.marpol.2016.07.026

Carruthers, T. R., & Hordyk, A. R. (2018a). The Data-Limited Methods Toolkit (DLMtool): An R package for informing management of data-limited populations. *Methods in Ecology and Evolution*, *9*(12), 2388–2395. https://doi.org/10.1111/2041-210X.13081

Carruthers, T. R., & Hordyk, A. R. (2018b). Using management strategy evaluation to establish indicators of changing fisheries. *Canadian Journal of Fisheries and Aquatic Sciences*, 1–16, https://doi.org/10.1139/cjfas-2018-0223

Carruthers, T. R., & Hordyk, A. R. (2019). *DLMtool: Data-limited methods toolkit*. Retrieved from https://cran.r-project.org/package=DLMtool

Carruthers, T. R., Punt, A. E., Walters, C. J., MacCall, A., McAllister, M. K., Dick, E. J., & Cope, J. (2014). Evaluating methods for setting catch limits in data-limited fisheries. *Fisheries Research*, *153*, 48–68. https://doi.org/10.1016/j.fishres.2013.12.014

Carvalho, F., Punt, A. E., Chang, Y. J., Maunder, M. N., & Piner, K. R. (2017). Can diagnostic tests help identify model misspecification in integrated stock assessments? *Fisheries Research*, *192*, 28–40. https://doi.org/10.1016/j.fishres.2016.09.018

Clark, W. (1999). Effects of an erroneous natural mortality rate on a simple age-structured stock assessment. *Canadian Journal of Fisheries and Aquatic Sciences*, *56*(10), 1721–1731. https://doi.org/10.1139/f99-085

Cochrane, K. L., Butterworth, D. S., De Oliveira, J. A. A., & Roel, B. A. (1998). Management procedures in a fishery based on highly variable stocks and with conflicting objectives: Experiences in the South African pelagic fishery. *Reviews in Fish Biology and Fisheries*, *8*, 177–214. https://doi.org/10.4161/hv.25031

Cooke, J. G. (1999). Improvement of fishery-management advice through simulation testing of harvest algorithms. *ICES Journal of Marine Science*, *56*(6), 797–810. https://doi.org/10.1006/jmsc.1999.0552

Deroba, J. J., & Schueller, A. M. (2013). Performance of stock assessments with misspecified age- and time-varying natural mortality. *Fisheries Research*, *146*, 27–40. https://doi.org/10.1016/j.fishres.2013.03.015

Forrest, R. E., McAllister, M. K., Dorn, M. W., Martell, S. J. D., & Stanley, R. D. (2010). Hierarchical Bayesian estimation of recruitment parameters and reference points for Pacific rockfishes (*Sebastes* spp.) under alternative assumptions about the stock–recruit function. *Canadian Journal of Fisheries and Aquatic Sciences*, *67*(10), 1611–1634. https://doi.org/10.1139/f10-077

Goodyear, C. P. (1977). Assessing the impact of power plant mortality on the compensatory reserve of fish populations. In W. Van Winkle (Ed.), *Proceedings of the conference on assessing the effects of power-plant-induced mortality on fish populations* (pp. 186–195). Pergamon. https://doi.org/10.1016/B978-0-08-021950-9.50021-1

Hamel, O. S. (2015). A method for calculating a meta-analytical prior for the natural mortality rate using multiple life history correlates. *ICES Journal of Marine Science*, *72*(1), 62–69.

Harley, S. J., Myers, R. A., & Dunn, A. (2001). Is catch-per-unit-effort proportional to abundance? *Canadian Journal of Fisheries and Aquatic Sciences*, *58*(9), 1760–1772. https://doi.org/10.1139/f01-112

Hilborn, R. (2003). The state of the art of stock assessment: Where are we going. *Scientia Marina*, *67*(Suppl. 1), 15–20.

Hilborn, R. (2010). Pretty Good Yield and exploited fishes. *Marine Policy*, *34*(1), 193–196. https://doi.org/10.1016/j.marpol.2009.04.013

Hilborn, R., & Walters, C. J. (1992). *Quantitative fisheries stock assessment: Choice, dynamics, and uncertainty*. New York, NY: Chapman and Hall.

Hoenig, J. M. (1983). Empirical use of longevity data to estimate mortality-rates. *Fishery Bulletin*, *81*, 898–903. https://doi.org/10.2307/1940298

Hurtado-Ferro, F., Szuwalski, C. S., Valero, J. L., Anderson, S. C., Cunningham, C. J., Johnson, K. F., … Punt, A. E. (2015). Looking in the rear-view mirror: Bias and retrospective patterns in integrated, age-structured stock assessment models. *ICES Journal of Marine Science: Journal Du Conseil*, *72*(1), 99–110. https://doi.org/10.1093/icesjms/fsu223

Huynh, Q., Hordyk, A. R., & Carruthers, T. R. (2018). *MSEtool: Management strategy evaluation toolkit*. Retrieved from https://cran.r-project.org/package=MSEtool

ICCAT (2017). Report of the 2017 ICCAT bluefin stock assessment meeting. Madrid, Spain.

Kell, L. T., Kimoto, A., & Kitakado, T. (2016). Evaluation of the prediction skill of stock assessment using hindcasting. *Fisheries Research*, *183*, 119–127. https://doi.org/10.1016/j.fishres.2016.05.017

Kenchington, T. J. (2014). Natural mortality estimators for information-limited fisheries. *Fish and Fisheries*, *15*(4), 533–562. https://doi.org/10.1111/faf.12027

Kristensen, K., Nielson, A., Berg, C. W., Skaug, H., & Bell, B. M. (2016). TMB: Automatic differentiation and laplace approximation. *Journal of Statistical Software*, *70*(5), 1–21. https://doi.org/10.18637/jss.v070.i0

Lee, H. H., Maunder, M. N., Piner, K. R., & Methot, R. D. (2012). Can steepness of the stock-recruitment relationship be estimated in fishery stock assessment models? *Fisheries Research*, *125–126*, 254–261. https://doi.org/10.1016/j.fishres.2012.03.001

Legault, C. M. (2009). *Report of the retrospective working group, January 14-16, 2008*. Woods Hole, Massachusetts. Retrieved from https://www.nefsc.noaa.gov/publications/crd/crd0901/

Legault, C. M., & Palmer, M. C. (2016). In what direction should the fishing mortality target change when natural mortality increases within an assessment? *Canadian Journal of Fisheries and Aquatic Sciences*, *73*(May 2015), 349–357. https://doi.org/10.1139/cjfas-2015-0232

Lowerre-Barbieri, S. K., Brown-Peterson, N. J., Murua, H., Tomkiewicz, J., Wyanski, D. M., & Saborido-Rey, F. (2011). Emerging issues and methodological advances in fisheries reproductive biology. *Marine and Coastal Fisheries*, *3*(1), 32–51. https://doi.org/10.1080/19425120.2011.555725

Lowerre-Barbieri, S. K., Ganias, K., Saborido-Rey, F., Murua, H., & Hunter, J. R. (2011). Reproductive timing in marine fishes: Variability, temporal scales, and methods. *Marine and Coastal Fisheries*, *3*(1), 71–91. https://doi.org/10.1080/19425120.2011.556932

MacCall, A. D. (2009). Depletion-corrected average catch: A simple formula for estimating sustainable yields in data-poor situations. *ICES Journal of Marine Science*, *66*(10), 2267–2271. https://doi.org/10.1093/icesjms/fsp209

Mangel, M., MacCall, A. D., Brodziak, J., Dick, E. J., Forrest, R. E., Pourzand, R., … Rose, K. (2013). A perspective on steepness, reference points, and stock assessment. *Canadian Journal of Fisheries and Aquatic Sciences*, *70*(6), 930–940. https://doi.org/10.1139/cjfas-2012-0372

Maunder, M. N., & Piner, K. R. (2015). Contemporary fisheries stock assessment: Many issues still remain. *ICES Journal of Marine Science*, *72*(1), 7–18. https://doi.org/10.1093/icesjms/fsu015

Maunder, M. N., & Punt, A. E. (2013). A review of integrated analysis in fisheries stock assessment. *Fisheries Research*, *142*, 61–74. https://doi.org/10.1016/j.fishres.2012.07.025

Omori, K. L., Hoenig, J. M., Luehring, M. A., & Baier-Lockhart, K. (2016). Effects of underestimating catch and effort on surplus production models. *Fisheries Research*, *183*, 138–145. https://doi.org/10.1016/j.fishres.2016.05.021

Pauly, D. (1980). On the interrelationships between natural mortality, growth parameters, and mean environmental temperature in 175 fish stocks. *ICES Journal of Marine Science*, *39*(2), 175–192. https://doi.org/10.1093/icesjms/39.2.175

Pauly, D., & Zeller, D. (2016). Catch reconstructions reveal that global marine fisheries catches are higher than reported and declining. *Nature Communications*, *7*, 1–9. https://doi.org/10.1038/ncomms10244

Punt, A. E., Butterworth, D. S., de Moor, C. L., De Oliveira, J. A. A., & Haddon, M. (2016). Management strategy evaluation: Best practices. *Fish and Fisheries*, *17*(2), 303–334. https://doi.org/10.1111/faf.12104

Punt, A. E., & Dorn, M. (2014). Comparisons of meta-analytic methods for deriving a probability distribution for the steepness of the stock-recruitment relationship. *Fisheries Research*, *149*, 43–54. https://doi.org/10.1016/j.fishres.2013.09.015

Punt, A. E., & Hilborn, R. (1997). Fisheries stock assessment and decision analysis: The Bayesian approach. *Reviews in Fish Biology and Fisheries*, *7*, 35–63. https://doi.org/10.1023/A:1018419207494

Punt, A. E., Hurtado-Ferro, F., & Whitten, A. R. (2014). Model selection for selectivity in fisheries stock assessments. *Fisheries Research*, *158*, 124–134. https://doi.org/10.1016/j.fishres.2013.06.003

R Core Team (2018). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.r-project.org/

Rice, J. C., & Rochet, M.-J. (2009). Simulation-based management strategy evaluation: Ignorance disguised as mathematics? *ICES Journal of Marine Science*, *66*(4), 754–762. https://doi.org/10.1093/icesjms/fsp023

Rudd, M. B., & Branch, T. A. (2017). Does unreported catch lead to overfishing? *Fish and Fisheries*, *18*(2), 313–323. https://doi.org/10.1111/faf.12181

Sampson, D. B., Hamel, O. S., Fisheries, N., Bosley, K., Budrick, J., Cronin-fine, L., …Whitman, A. D. (2017). 2017 Assessment update for the US West Coast stock of Arrowtooth Flounder. Portland, Oregon.

SEDAR (2005). Stock assessment report of SEDAR 7, Gulf of Mexico red snapper. Southeast data, assessment, and review.

SEDAR (2008). SEDAR 17: Stock assessment report 1: South Atlantic red snapper. Southeast data, assessment, and review.

SEDAR (2010). SEDAR 24: Stock assessment report: South Atlantic red snapper. Southeast data, assessment, and review.

SEDAR (2015). SEDAR 31 update stock assessment report: Stock assessment of red snapper in the Gulf of Mexico 1872–2013 with provisional 2014 landings.

SEDAR (2017). SEDAR 41: Stock assessment report – Revision 1: South Atlantic red snapper. Southeast data, assessment, and review.

Stewart, I. J., & Martell, S. J. D. (2014). A historical review of selectivity approaches and retrospective patterns in the Pacific halibut stock assessment. *Fisheries Research*, *158*, 40–49. https://doi.org/10.1016/j.fishres.2013.09.012

Then, A. Y., Hoenig, J. M., Hall, N. G., & Hewitt, D. A. (2015). Evaluating the predictive performance of empirical estimators of natural mortality rate using information on over 200 fish species. *ICES Journal of Marine Science*, *72*(1), 82–92. https://doi.org/10.1093/icesjms/fss153

Thompson, G. G. (1994). Confounding of gear selectivity and the natural mortality rate in cases where the former is a nonmonotone function of age. *Canadian Journal of Fisheries and Aquatic Sciences*, *51*(12), 2654–2664. https://doi.org/10.1139/f94-265

Tuck, G. N. (Ed.) (2016). *Stock assessment for the Southern and Eastern Scalefish and Shark Fishery: 2015. Part I*. Hobart, Australia: Australian Fisheries Management Authority and CSIRO Oceans and Atmosphere.

Van Beveren, E., Duplisea, D., Castonguay, M., Doniol-Valcroze, T., Plourde, S., & Cadigan, N. (2017). How catch underreporting can bias stock assessment of and advice for northwest Atlantic mackerel and a possible resolution using censored catch. *Fisheries Research*, *194*(May), 146–154. https://doi.org/10.1016/j.fishres.2017.05.015

Walters, C. J., & Martell, S. J. D. (2004). *Fisheries ecology and management*. Princeton, NY: Princeton University Press.

Walters, C. J., Martell, S. J. D., & Korman, J. (2006). A stochastic approach to stock reduction analysis. *Canadian Journal of Fisheries and Aquatic Sciences*, *63*(1), 212–223. https://doi.org/10.1139/f05-213

Wetzel, C. R., Cronin-Fine, L., & Johnson, K. F. (2017). *Status of Pacific ocean perch (*Sebastes alutus*) along the US west coast in 2017*. Seattle, Washington, USA. Retrieved from https://www.pcouncil.org/wp-content/uploads/2018/01/PacificOceanPerch2017_Assessment_Final.pdf

Williams, E. H. (2002). The effects of unaccounted discards and misspecified natural mortality on harvest policies based on estimates of spawners per recruit. *North American Journal of Fisheries Management*, *22*, 311–325. https://doi.org/10.1577/1548-8675(2002)022<0311:-TEOUDA>2.0.CO;2

Zhu, J., Chen, Y., Dai, X., Harley, S. J., Hoyle, S. D., Maunder, M. N., & Aires-da-Silva, A. M. (2012). Implications of uncertainty in the spawner-recruitment relationship for fisheries management: An illustration using bigeye tuna (*Thunnus obesus*) in the eastern Pacific Ocean. *Fisheries Research*, *119–120*, 89–93. https://doi.org/10.1016/j.fishres.2011.12.008

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Hordyk AR, Huynh QC, Carruthers TR. Misspecification in stock assessments: Common uncertainties and asymmetric risks. *Fish Fish*. 2019;20:888–902. https://doi.org/10.1111/faf.12382