Fisheries and Oceans Canada
Pêches et Océans Canada

Ecosystems and Oceans Science
Sciences des écosystèmes et des océans

Canadian Science Advisory Secretariat (CSAS)

Research Document 2024/029

National Capital Region

Technical Considerations for Stock Status and Limit Reference Points under the Fish Stocks Provisions

Tim J. Barrett[1], Julie R. Marentette[2], Robyn E. Forrest[3], Sean C. Anderson[3], Carrie A. Holt[3], Danny W. Ings[2], Mary E. Thiess[2]

[1]St. Andrews Biological Station
Fisheries and Oceans Canada
125 Marine Science Drive
St. Andrews, NB, E5B 0E4

[2]Ecosystems and Oceans Science Sector
Fisheries and Oceans Canada
200 Kent Street
Ottawa, ON, K1A 0E6

[3]Pacific Biological Station
Fisheries and Oceans Canada
3190 Hammond Bay Road
Nanaimo, BC, V9T 6N7

April 2024

Canada

Foreword

This series documents the scientific basis for the evaluation of aquatic resources and ecosystems in Canada. As such, it addresses the issues of the day in the time frames required and the documents it contains are not intended as definitive statements on the subjects addressed but rather as progress reports on ongoing investigations.

Published by:

Fisheries and Oceans Canada
Canadian Science Advisory Secretariat
200 Kent Street
Ottawa ON K1A 0E6

http://www.dfo-mpo.gc.ca/csas-sccs/
csas-sccs@dfo-mpo.gc.ca

ISSN 1919-5044
ISBN 978-0-660-71266-6 Cat. No. Fs70-5/2024-029E-PDF

Correct citation for this publication:

Barrett, T.J., Marentette, J.R., Forrest, R.E., Anderson, S.C., Holt, C.A., Ings, D.W., and Thiess, M.E. 2024. Technical Considerations for Stock Status and Limit Reference Points under the Fish Stocks Provisions. DFO Can. Sci. Advis. Sec. Res. Doc. 2024/029. v + 57 p.

Aussi disponible en français :

Barrett, T.J., Marentette, J.R., Forrest, R.E., Anderson, S.C., Holt, C.A., Ings, D.W., et Thiess, M.E. 2022. Considérations techniques relatives à l’état des stocks et aux points de référence limites dans le cadre des dispositions relatives aux stocks de poissons. Secr. can. des avis sci. du MPO. Doc de rech. 2024/029. v + 66 p.

TABLE OF CONTENTS

ABSTRACT

Revisions to Canada's *Fisheries Act* have resulted in a need for a single limit reference point (LRP) and metric of stock status for major fish stocks prescribed by regulation. The Science Sector has identified a need to provide guidance to estimate LRPs and stock status for scenarios that presently do not meet the "one stock, one LRP, one status" requirement, and a more general need for guidance on methods to estimate and report both LRP and stock status across a spectrum of data and knowledge availability and quality. To inform this guidance, we provide a review of literature and approaches to defining LRPs, describe technical considerations for choosing from various approaches for estimating LRPs and indicators of stock status across the data spectrum, and provide technical considerations and guidance for estimating a single LRP and metric of stock status in cases where either a single assessment model or multiple models are applied. We review methods to estimate B_{MSY} and B_0; theoretical, historical, and empirical proxies for these indicators; and some generic "rules of thumb" for other common LRPs used in Canada. We also provide examples of less common indicators, LRPs, and stock status estimation methods that may be applicable across the data spectrum, and review approaches with which to address volatility in stock status indicators. We provide operational and technical considerations as a basis from which to select or reject various candidate indicators and LRPs as well as options and considerations for reporting a single status per stock in an assessment or stock status update, and across advice and management frameworks.

1. INTRODUCTION

1.1. CONTEXT

Canada's *Fisheries Act* (R.S.C., 1985, c. F-14) was revised on June 21, 2019, resulting in new Considerations and Fish Stocks Provisions (FSPs) that relate to the management of fisheries. The requirements of the FSPs are being interpreted through the lens of the Fisheries and Oceans Canada's (DFO) Sustainable Fisheries Framework suite of policies, in particular the *Fishery decision-making framework incorporating the precautionary approach* (PA Policy, DFO 2009). The FSPs came into force in April 2022 with the prescription of the first batch of *major fish stocks* and associated regulations for rebuilding plan requirements.

The FSPs identify objectives for conservation in light of sustainable use. These include objectives to maintain stocks at or above the "level necessary to promote sustainability of the stock", s 6.1(1); or "above [the limit reference point or LRP]", s 6.1(2); "rebuild [stocks above LRP]", s 6.2(1); as well as "minimizing further [stock] decline" and mitigating "adverse socio-economic or cultural impacts", s 6.2(2)). The FSPs also reference required management actions (setting LRPs, implementing measures, and developing and implementing rebuilding plans). The FSPs are understood to require advice pertaining to:

- A single LRP for each candidate or prescribed major fish stock (the "one-stock-one-LRP" requirement);
- A single stock status relative to the LRP (part of the objective for s 6.1(2) and when breached, a rebuilding plan is triggered under s 6.2); and
- Supporting the prescription of major fish stocks that are of one species, can be defined geographically or by management area, and for which there is a single LRP/stock status.

In Canada, the LRP is intended to help operationalize an objective to avoid *serious harm* to fish stocks, a central objective of the PA Policy (DFO 2009). In the PA Policy, the LRP serves multiple roles. As a limit, it is part of fisheries management objectives. As part of stock status metrics, the LRP defines the Critical Zone. The LRP further serves as an inflection point for the removal reference (limit fishing mortality rate), and often as an operational control point for management measures. Lastly, the LRP now has a legislated role in triggering the need for rebuilding plans when breached. Fisheries managers have a leading role for most aspects of fisheries management objectives and therefore the PA Policy (including most reference points, risk tolerance and timeframes over which objectives are evaluated, and measures to achieve the objectives), the responsibility for setting LRPs (and the estimation of stock status) lies with DFO's Ecosystems and Oceans Science Sector (Science Sector; DFO 2009).

1.2. GOALS OF THIS PAPER

This research document is the second of a series of three working papers presented during the June 21-29, 2022 National Advisory Process meeting entitled "Science Advice on Guidance for Limit Reference Points under the Fish Stocks Provisions" (DFO 2023a). The first working paper contains a review of Canadian approaches to operationalizing LRPs and stock status, and a literature review of existing Canadian and international guidance for LRP and stock status estimation. A consolidated definition of serious harm is provided in the first paper and a set of six best-practice principles concerning indicators, LRPs and stock status metrics are identified

(Marentette et al. unpublished working paper)[1]. The third paper provides guidance and recommendations on identifying a single LRP and status where there is a complex spatial and temporal relationship between stock structure and associated biological processes, stock and fishery monitoring data, and management measures applied to the stock (Ings et al. unpublished working paper) [2]. To support the process of choosing LRPs and estimating stock status, Marentette et al. (unpublished working paper) suggested technical guidance should focus on:

- Different methods with which to estimate biomass at maximum sustainable yield (B_{MSY}) and unfished biomass (B_0) suitable across the data spectrum;
- Theoretical, historical, and empirical proxies for model-based indicators and for B_{MSY} and B_0 or carrying capacity (K);
- Generic "rules of thumb" for other common LRPs used in Canada (e.g., proportions of B_0 or K, the minimum biomass that produced the recruitment that lead to stock recovery ($\mathrm{B}_{recover}$), recruitment thresholds);
- Examples of less common indicators, LRPs, and stock status estimation methods that may be applicable across the data spectrum (including composites, "traffic light" approaches, and expert judgement);
- Different approaches with which to address volatility in stock status indicators;
- Use of basic life history information and meta-analyses (including extrapolation from other stocks) to inform LRP selection;
- Operational and technical considerations as a basis from which to select or reject various candidate indicators and LRPs, e.g.,
    - pros and cons,
    - diagnostics to evaluate reliability of estimation, including sensitivity analyses, and
    - considerations for evaluating the plausibility of candidate LRPs given the history of the stock, past evidence of serious harm, other biological information or analogous stocks;
- Options and considerations for reporting a single status per stock, within and between assessments, and across advice and management frameworks.

These needs were addressed in this paper by:

- Providing a review of literature and approaches to defining LRPs for practitioners;
- Describing technical considerations for choosing from among various approaches for estimating LRPs and indicators of stock status across the data spectrum, including situations of model uncertainty and parameter uncertainty; and
- Providing technical considerations and guidance for estimating a single LRP and metric of stock status in cases where either a single assessment model or multiple models are applied.

---

[1] Marentette, J.R., Barrett, T.J., Cogliati, K.M., Ings, D.W., Ladell, J., and Thiess, M.E. Operationalizing thresholds to serious harm: existing guidance and contemporary Canadian practices. Unpublished working paper.

[2] Ings, D.W., Marentette, J.R., Thiess, M.E., and Barrett T.J. Considerations for stock structure and management scale under the Fish Stocks Provisions. Unpublished working paper.

To achieve these goals, we reviewed multiple methods of estimating LRPs, identified how the methods relate to serious harm, and described the relationships among the different methods based on the assumed resilience of the stock. The selection of indicators of stock status and selection of an LRP will be stock-specific in most cases. Therefore, this paper offers general guidance to support practitioners, instead of prescribed guidance for which there would be numerous exceptions. This general guidance is intended to help practitioners identify methods that are consistent with international best practices and consistent with the best practice principles for indicators, LRPs, and stock status defined by Marentette et al. (unpublished working paper).

2. SELECTION OF LIMIT REFERENCE POINTS

LRPs in Canadian policy are typically defined in terms of biomass, particularly spawning stock biomass (SSB), abundance, or a proxy for these, but other metrics that represent the reproductive capacity (e.g., total egg production) of the stock may be used. The indicator(s) chosen should be the one(s) that best represents the reproductive capacity of the stock, depending on the type of analytical assessment (e.g., age-structured vs. surplus production model) and data availability. The biomass indicator in this paper could therefore represent SSB, female SSB, total biomass, or vulnerable biomass.

LRPs are commonly defined in terms of a proportion of unfished biomass (B_0) or biomass at maximum sustainable yield (B_{MSY}) or a proxy for these reference points. Generic policy guidance on LRPs in terms of a proportion of B_0 and/or B_{MSY} is provided by several jurisdictions, in the absence of which it is generally not clear which reference point and what proportion should be used. The relationship between B_0, B_{MSY}, and equilibrium biomass at $F_{X\%SPR}$ (fishing mortality rate that results in a spawning potential ratio of X%) depends strongly on the resilience of the stock or steepness (h) of the stock-recruitment relationship (SRR) (Section 2.1.4) and it is important to consider this relationship when a proportion of one reference point is used as a proxy for the other. Understanding how these metrics of stock status are influenced by model assumptions and life-history characteristics can help guide identification of a candidate LRP. An evaluation of multiple candidate LRPs estimated using different metrics can provide confidence in selecting a metric when estimates agree, but can also identify potential risks when metrics don't agree. For example, 0.4 B_{MSY} could occur at a very low depletion (biomass relative to B_0) where other forms of serious harm could occur. The pros and cons of different candidate LRPs are listed in Table 1 and some reasons why various methods have been chosen in Canada are provided in Marentette et al. (unpublished working paper).

Some alternative LRPs that are not listed in Table 1 include unique LRPs for semelparous stocks and Atlantic salmon stocks (Sections 2.8 and 2.10) and alternative approaches to estimating stock status using a traffic light approach (Section 2.7) or expert judgement (Section 2.12). The traffic light approach has the advantage of including multiple indicators and has been used in a framework consistent with the PA Policy for Pacific salmon (DFO 2022a). Expert judgement may be used for determining stock status; however, it may be difficult in some situations to formally define the LRP and quantify uncertainty. Stock status determination by expert judgement can be biased based on the level of expertise of the experts. A weight-of-evidence approach based on expert judgement can be used to arrive at a status determination using multiple indicators based on cumulative weight-of-evidence, in a transparent and reproducible way.

Table 1. Candidate LRPs, their link(s) to serious harm, and some pros and cons. h = steepness of stock-recruitment relationship (SRR). M = natural mortality rate. F_{MSY} = fishing mortality rate at MSY. $F_{X\%SPR}$ = fishing mortality rate at X% spawning potential ratio. R_{max} = maximum recruitment. *B_{rep} = long-term equilibrium SSB that results from fishing at the replacement fishing mortality rate (F_{rep}). B_{loss} = lowest observed biomass.*

LRP:	**Proportion of B_{MSY} (Section 2.1)**
Link(s) to Serious Harm	Loss of surplus production; Proxy for relative depletion (proportion of B_0) and thus for recruitment overfishing.
Pros	DFO provisional PA Policy default LRP (0.4 B_{MSY}).
Cons	May be difficult to estimate; Sensitive to uncertainty in selectivity, *M*, and *h* of the SRR; May correspond to very low biomass where stock dynamics may be less well known (e.g., validity of compensatory assumptions) or other sources of serious harm may become important (e.g., Allee effects).
LRP:	**Proportion of B_0 (Section 2.1)**
Link(s) to Serious Harm	Proxy for recruitment overfishing; Proxy for Allee effects, Proportions of B_0 may be a proxy for B_{MSY}.
Pros	B_0 may be more reliably estimated than B_{MSY}.
Cons	No explicit DFO PA Policy provisional default for the proportion to choose; B_0 may be difficult to estimate; Sensitive to uncertainty in model assumptions (e.g., *M*); If used as a proxy for 0.4 B_{MSY}, the ratio B_{MSY} / B_0 depends strongly on the productivity of the stock (e.g., *h*, which is often poorly estimated), as well as the relationship between maturity-at-age and fishery selectivity-at-age.
LRP:	**Proportion of Theoretical Proxies for B_{MSY} (Section 2.2)**
Link(s) to Serious Harm	Loss of reproductive potential; Proxy for B_{MSY}.
Pros	Requires fewer assumptions and data (e.g., no SRR); $F_{X\%SPR}$ (with X = 40 commonly used) and others can be a proxy for F_{MSY}, and therefore can be used to estimate a proxy for B_{MSY}.
Cons	X depends on stock productivity; Once X is selected, $F_{X\%SPR}$ is sensitive to *M*; A per-recruit estimate does not account for lower recruitment at low biomass; An estimate of equilibrium recruitment is needed to estimate the equilibrium biomass at $F_{X\%SPR}$.
LRP:	**Biomass at X% R_{max} or other thresholds to impaired recruitment (Section 2.3)**
Link(s) to Serious Harm	Loss of recruitment; R_{max} estimated from a Ricker SRR or B_{rep} (see Section 2.3) is sometimes used as a proxy for B_{MSY}.
Pros	Easy to interpret; Stock specific thresholds are possible.
Cons	Dependent on SRR (including data on recruitment at low stock sizes); May occur at a very low level of depletion for stocks with high *h*.

Table 1 (continued). Candidate LRPs, their link(s) to serious harm, and some pros and cons.

LRP:	**$B_{recover}$ (including B_{loss}, etc.) (Section 2.4)**
Link(s) to Serious Harm	Proxy for recruitment overfishing, reflecting high uncertainty in population dynamics at low stock sizes.
Pros	Easy to understand and communicate; Not influenced as strongly by model assumptions; Recommended for stocks with occasional large year classes (spasmodic recruitment).
Cons	Values may vary widely among stocks (they may not “scale” with stock size or life history); Assumption of possible recovery in future depends on prevailing conditions in relation to the reference period; Recovery must also be defined and no consistent practice has emerged for what constitutes recovery. Variation in practices for selecting year(s) used in defining these reference points.
LRP:	**Historical Proxies for B_{MSY} or B_0 (Appendix B)**
Link(s) to Serious Harm	Often employed as proxies for other reference points such as B_{MSY} or B_0, although LRPs based on other thresholds to serious harm may be considered (e.g., agreed-upon undesirable states to avoid).
Pros	Easy to understand and communicate; Can be applied to data-limited stocks; A provisional proxy for B_{MSY} is 50% of the maximum population size as suggested by the PA Policy.
Cons	Relies on assumptions of the relationship between the indicator and stock attribute it represents; Wide variation in practice for selecting historical time periods, and thus may be variation in suitability for approximating B_{MSY} or B_0.
LRP:	**Empirical LRPs (Appendix B)**
Link(s) to Serious Harm	Often employed as proxies for other reference points such as B_{MSY}, B_0, or $B_{recover}$, although other thresholds to serious harm may be considered (e.g., agreed-upon undesirable states to avoid).
Pros	Easy to understand and communicate; Can be applied to data-limited stocks; Based on observable quantities that do not rely on model assumptions.
Cons	May be harder to link to desired management outcomes in some cases; Relies on assumptions of the relationship between the indicator and stock attribute it represents.

2.1. LRPS BASED ON B_0 AND B_{MSY}

2.1.1. Thresholds of Relative Depletion (Proportions of B_0)

B_0 can be defined as the mean long-term equilibrium biomass of the stock in the absence of fishing. In general, B_0 is assumed to be stationary over time (i.e., stock and prevailing environmental conditions are assumed stable over time). In some cases these assumptions may be violated (see Section 5.1). B_0 can also represent the theoretical carrying capacity (K) of the stock in surplus production models (e.g., Schaefer 1954; Section A.2). To avoid misinterpretation, it is important to define B_0 and the assumptions used to define it (e.g., the time period over it reflects the equilibrium biomass).

A proportion of B_0 (e.g., 0.2 B_0) is commonly adopted as an LRP, which may provide a reasonable proxy threshold for recruitment overfishing for productive stocks because there is often at least some reduction in per-capita recruitment below this threshold (Myers et al. 1994; Sainsbury 2008). A threshold of 0.2 B_0 is the default recommendation and minimum acceptable value for an LRP in the absence of a stock-specific choice under Australia's Commonwealth Fisheries Harvest Strategy Policy (DAWR 2018). Similarly, it is one default for the soft limit (a trigger for rebuilding plans) in New Zealand's Harvest Strategy (MF 2008). For less productive stocks, the corresponding threshold is often recommended to be higher, e.g., 0.3 B_0 (Musick 1999, Mace et al. 2002, Sainsbury 2008; Stenson et al. 2012; DAWR 2018). Sainsbury (2008) further recommends 0.3 B_0 as a best practice LRP for low productivity stocks unless there are high fluctuations in productivity, in which case 0.2 of the median long-term $B_{unfished}$ may be more appropriate, where he defines $B_{unfished}$ as a time-varying proxy for B_0 that does not assume equilibrium conditions.

2.1.2. Thresholds for Impaired Production (Proportions of B_{MSY})

The concept of MSY has been the global standard for sustainable fisheries for decades (Mace 2001; Punt and Smith 2001; Cadrin 2012) and is included in many international agreements (e.g., the Straddling Fish Stocks and Highly Migratory Fish Stocks Agreement, UN 1995; the European Union Common Fisheries Policy, EU 2013) and national laws (e.g., US Magnuson-Stevens Fishery Conservation and Management Act, eCFR 2021). Theoretically, MSY is the maximum long-term yield that the stock can produce, given constant life history and selectivity parameters. MSY occurs at an intermediate level of fishing mortality (F, ranging between $F = 0$ and the F that would cause extirpation of the stock) and at a biomass below K or B_0. Conceptually, with increasing long-term F, the abundance and mean age of a population decreases and the per-capita growth rate of the population increases as a result of reduced competition or similar effects as the biomass reduces from K to B_{MSY} (Sainsbury 2008). The increase in the per-capita growth rate (e.g., due to increased growth, survival, and reproductive success at lower population density) results in a surplus of biomass (i.e., surplus production) that can be harvested. Surplus production, like recruitment, is another direct measure of stock productivity (NAFO 2004) and persistent states of low or negative surplus production can be evidence of serious harm (Kronlund et al. 2018).

There appear to be at least three interpretations of what LRPs based on proportions of B_{MSY} are intended to represent:

1. a biomass level associated with an unacceptable reduction in surplus production;
2. a biomass level associated with less desirable yields; or
3. a proportion of B_{MSY} as a proxy for another proportion of B_0 and/or a proxy for recruitment overfishing.

LRPs based on proportions of B_{MSY} are common in Canada and elsewhere (Sainsbury 2008; DFO 2016, Marentette et al. unpublished working paper), familiar to managers and fishery interests (McKown et al. 2008), and often given in association with policy guidance on provisional default values (e.g., 0.4 B_{MSY} in Canada; DFO 2009). Some jurisdictions use 0.3 B_{MSY} as a default LRP for stocks where production is estimated (NAFO 2004, ICES 2021a), since that is the biomass associated with a 50% reduction in maximum surplus production, under the assumptions of a Schaefer (1954) production model (Figure 1). A 50% reduction in surplus production could arguably serve as a literal interpretation of impaired productivity. Limits of 0.5 B_{MSY} are often used in the US, but under legislative and policy frameworks where undesirable stock states are associated with an unacceptable reduction in ability to obtain optimum yield rather than serious harm to the stock (NAFO 2004). Loss of yield alone would not be consistent with an objective to avoid serious harm to the stock (Shelton and Rice 2002), but that does not necessarily exclude 0.5 B_{MSY} from (also) being consistent with a threshold to serious harm. In fact, some have noted that 0.5 B_{MSY} can occur at very low biomass, for example, 0.175 B_0 for species considered to be highly productive (Sainsbury 2008). Even for highly productive stocks, such biomass levels may be inconsistent with objectives of maintaining resilience and avoiding loss of genetic diversity (Sainsbury 2008) which can also be considered indicative of serious harm.

MSY-based reference points can be estimated in several ways (Punt et al. 2014; Appendix A). The simplest method of estimating MSY, and associated B_{MSY} and F_{MSY}, is to use a surplus production model such as the Schaefer model (Schaefer 1954; Figure 1; Section A.2). The minimum requirement for fitting a surplus production model is a time series of catch and an index of abundance or biomass that are used to estimate the parameters r and K, where r is the intrinsic rate of population growth and K is the population carrying capacity. Using an age-structured model, MSY reference points are estimated using the SRR, growth, maturity, M, and fishery selectivity data. In addition to catch and index data, age-structured models can also be fit to age and/or length composition data, which can inform estimates of M and SRR parameters. For certain SRRs, specifically with high values of the steepness (h) parameter of the SRR (h is the proportion of unexploited recruitment, R_0, produced at 0.2 B_0; Mace and Doonan 1988), B_{MSY} can occur at a very low level of depletion, where depletion is defined as biomass relative to B_0 (Mace 1994). For a symmetrical Schaefer production model, 0.5 B_{MSY} occurs at 75% of MSY (on the left-hand side of the yield curve; Figure 1) and DFO's PA Policy default LRP of 0.4 B_{MSY} is equivalent to 0.2 B_0 and occurs at 64% of MSY (Figure 1). When modelling with an asymmetric production model (e.g., Pella and Tomlinson 1969; Fletcher 1978) or an age-structured model SRR, these percentages will vary. For example, in an age-structured model the relationship between B_0 and B_{MSY} will be dependent on factors such as h (e.g., Figure 2), M, and the relationship of selectivity-at-age to maturity-at-age (see Section 2.1.3).

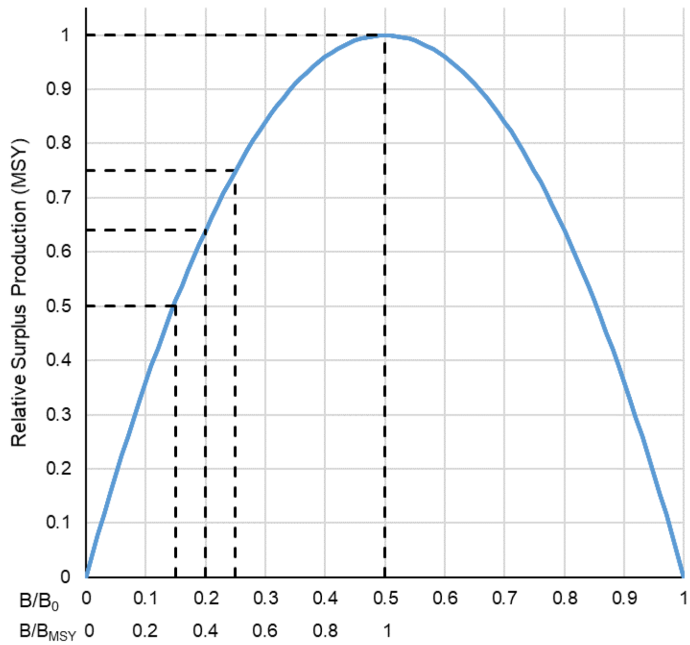

Figure 1. Plot of relative surplus production (maximum sustainable yield) vs. biomass as a ratio of B_0 (i.e., K) and B_{MSY} for a Schaefer production model. 0.3, 0.4, 0.5, and 1 B_{MSY} result in 50%, 64%, 75%, and 100% of MSY, respectively.

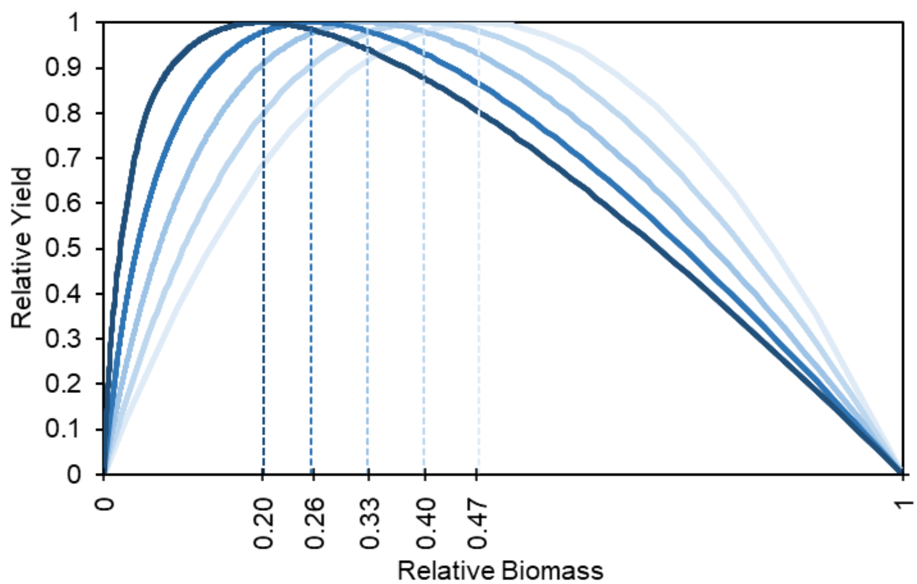

Figure 2. An example of changes in the shape of the relative yield vs. relative biomass curve and the position of B_{MSY} relative to B_0 (labels on x-axis) with changes in steepness (h) for a theoretical small pelagic stock with h = 0.95, 0.8, 0.6, 0.4, 0.25 (from dark to light). B_{MSY} relative to B_0 is identified on the x-axis for each curve.

2.1.3. Selecting Proportions of B_0 and B_{MSY} as LRPs

Some jurisdictions define generic policy guidance for LRPs first in terms of proportions of B_0, with proportions of B_{MSY} recommended under certain conditions. The LRP in Australian fisheries policy is defined as the biomass level where risk to the stock in terms of recruitment impairment is unacceptably high and must be set at or above 0.2 B_0 (DAWR 2018) under the assumption that 0.2 B_0 is a suitable threshold that avoids recruitment overfishing for productive stocks (Sainsbury 2008). For less productive stocks, Australia's harvest strategy policy and Sainsbury (2008) both suggest 0.3 B_0. An LRP of 0.5 B_{MSY} is only considered a potential alternative LRP in Australia's harvest strategy policy when B_{MSY} is reliably estimated and B_{MSY} is also above 0.4 B_0 (DAWR 2018; see Section 2.1.4). New Zealand's harvest policy recommends the use of either 0.2 B_0 or 0.5 B_{MSY}, whichever is higher, as a soft limit in their harvest strategies (and half of each again as a hard limit, below which closure of fisheries are considered). The Marine Stewardship Council recommends a default point of reproductive impairment of 0.2 B_0 = 0.5 B_{MSY} for stocks with "average productivity", in the absence of a stock-specific choice by making the assumption that B_{MSY} occurs at 0.4 B_0 (Agnew et al. 2014), but elaborates with more complex recommendations when more information is available. If B_{MSY} is greater than 0.4 B_0, then 0.5 B_{MSY} is advised, and if B_{MSY} is less than 0.4 B_0, then 0.2 B_0 is preferred (unless B_{MSY} < 0.27 B_0, in which case 0.75 B_{MSY} is recommended; MSC 2018).

Some jurisdictions only provide default policy guidance for limits based on B_{MSY}. The US guidance for minimum stock size thresholds is a minimum of 0.5 B_{MSY} (eCFR 2021). Both the Northwest Atlantic Fisheries Organization (NAFO) and the International Council for the Exploration of the Sea (ICES) recommend 0.3 B_{MSY} = 0.15 B_0 as LRPs for stocks assessed with surplus production models based on a loss of 50% of maximum production (NAFO 2004, ICES 2021a), although different choices are supported for stocks assessed with more complex assessment models. The choice of 0.4 B_{MSY} as a provisional default LRP for Canada was considered consistent with international best practices at the time (DFO 2009). If it is assumed that B_{MSY} = 0.5 B_0 (i.e., as for a Schaefer production model; DFO 2016), then 0.4 B_{MSY} would be conceptually equivalent to 0.2 B_0 (Figure 1).

2.1.4. The Relationship Between B_0 and B_{MSY}

A proportion of B_0 is sometimes used as a proxy for B_{MSY} when B_{MSY} cannot be directly or reliably estimated. The relationship between B_{MSY} and B_0 is determined by the shape of the yield vs. biomass plot (e.g., Figure 2) and this shape depends on the resilience of the stock to fishing pressure. When a proportion of B_0 is used as a proxy for $B_{MSY,}$ an implicit assumption about the productivity of the stock is made. Resilience of a stock to fishing is a function of the stock's productivity. In general, stocks with low productivity (e.g., high age-at-maturity, long-lived, slow growth rate) tend to be less resilient to fishing pressure (MF 2011). In a simple surplus production model (Schaefer 1954), resilience is represented in the intrinsic population growth parameter, *r*. In more complex models, the components of productivity include the growth rate of individuals, *M*, and *h*, which measure the degree of compensatory density-dependence in juvenile survival. Resilience to fishing is also partly determined by the relationship of fishery selectivity-at-age to maturity-at-age, where stocks will be more resilient if there is a portion of the mature population that is invulnerable to fishing (Myers and Mertz 1998).

Under a Beverton-Holt (BH) SRR, *h* is bounded between 0.2 and 1. This can result in values of B_{MSY}/B_0 ranging from <0.1 at very high (>0.95) *h* to 0.5 as *h* approaches 0.2. For age- or size-structured models, B_{MSY}/B_0 is strongly influenced by *h*, but is also influenced by *M*, maturity, selectivity, and weight-at-age (Figure 3). Punt et al. (2008) illustrated the variability in the relationship between B_{MSY}/B_0 and *h* among four groundfish stocks. Their estimates of B_{MSY}/B_0 of ~0.25 and ~0.35 corresponded to *h* values of ~0.8 to 0.9 and ~0.5 to 0.6, respectively (Punt et al. 2008). Punt et al. (2014) showed that B_{MSY}/B_0 was sensitive to maturity-at-age and selectivity-at-age for two groundfish species where B_{MSY}/B_0 can range from ~0.2 to ~0.4 for high *h* depending on the fecundity of unselected ages. These studies demonstrate the importance of understanding the influence of fishery and biological parameters on the MSY reference point estimates. Steepness is generally poorly estimated in SRRs and setting priors for *h* is common practice. Haltuch et al. (2008) showed that B_0 and current depletion could be estimated more reliably than B_{MSY}.

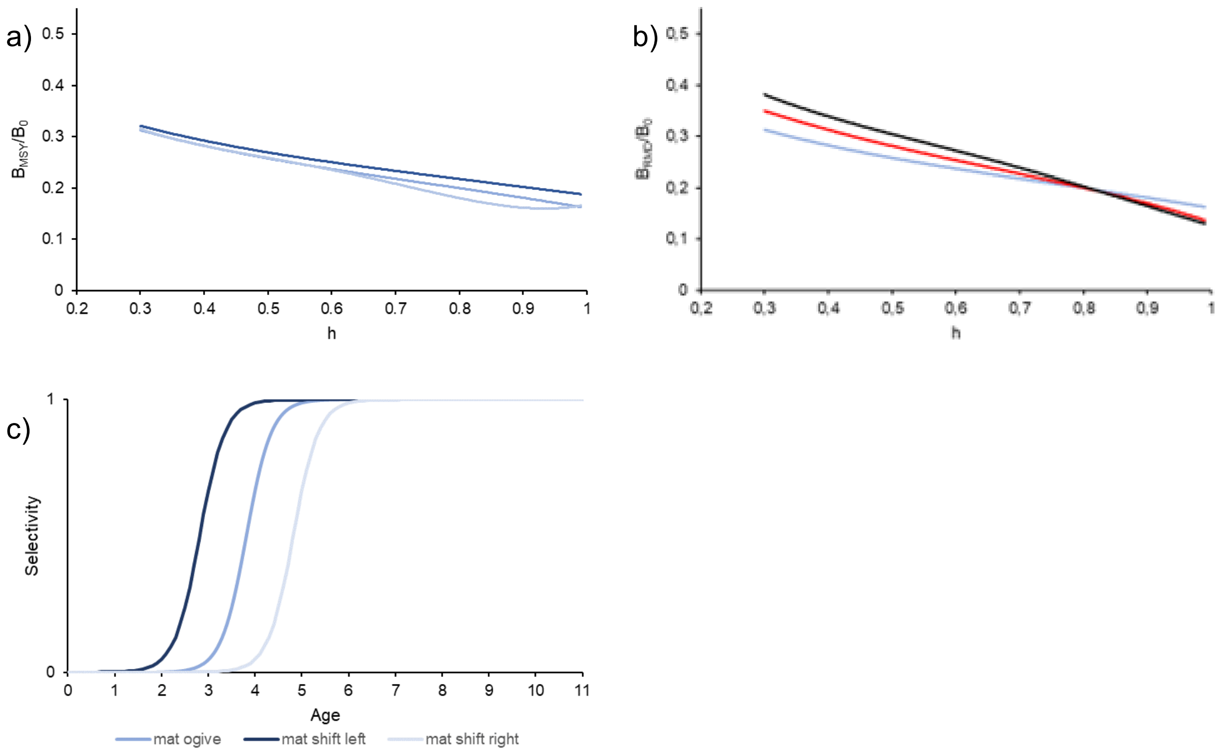

Figure 3. Relationship between B_{MSY}/B_0 and steepness (h) of the Beverton-Holt stock recruitment relationship for a small pelagic stock with a) M = 0.2 and three selectivity patterns: selectivity = maturity (mat) ogive, maturity ogive shifted left by one year, and maturity ogive shifted right by one age and b) M = 0.2, 0.3 (red), 0.4 (black) with selectivity = maturity ogive. Selectivity curves plotted in c).

Various proportions in the range of 0.3 to 0.6 of B_0 have been used to approximate B_{MSY} where the true ratio of B_{MSY}/B_0 is unknown, with higher proportions of B_0 used for less resilient species (Restrepo et al. 1998, Gabriel and Mace 1999). A meta-analysis of 147 stocks estimated a mean B_{MSY}/B_0 of 0.40, but a wide range of mean ratios (0.26 to 0.46) among taxonomic groups (Thorson et al. 2012). An investigation by Punt et al. (2014) across a range of uncertainties, revealed that B_{MSY} proxies in the range of 0.35 to 0.4 B_0 minimized the potential loss in yield compared to that if B_{MSY} was known exactly.

In Australia, it is recommended to consider 0.4 B_0 as the B_{MSY} proxy (DAWR 2018) and the Marine Stewardship Council similarly advises this default assumption for stocks with average productivity. A default policy equivalency of soft limits of 0.2 B_0 with 0.5 B_{MSY} in New Zealand also implies a default policy assumption that B_{MSY} = 0.4 B_0 (MF 2008), although recommended proxies for B_{MSY} also range from 0.25 to over 0.45 B_0 depending on stock productivity in accompanying operational guidelines (MF 2011; Table 2). The productivity classification used in MF (2011) and shown in Table 2 was determined by various life-history parameters, using categories defined by FAO (2001).

The implications of using the B_{MSY}/B_0 ratios in Table 2 in conjunction with the PA Policy default LRP of 0.4 B_{MSY} is a range of LRPs from 0.1 B_0 to 0.16 B_0 for high to low productivity stocks. These depletion levels may be inconsistent with objectives of avoiding serious harm (Section 2.1.2). In a recent evaluation of ICES stocks (ICES 2022), 50% of the 69 ICES Category 1 stocks had B_{lim} estimates (i.e., biomass-based LRPs) of less than 0.1 B_0 and a workshop recommendation was made to consider setting LRPs at a *minimum* of 0.1 B_0, which is consistent with the hard limit defined in New Zealand's harvest policy (MF 2008). Setting such a minimum policy limit of 0.2 B_0 (DAWR 2018) or 0.1 B_0 (ICES 2022) as an LRP, or setting minimum assumed relationships between B_0 and B_{MSY} (MF 2008), could avoid states of serious harm resulting from very low depletion.

DFO does not have default policy guidance for reference points based on B_0. Given the prevalence of LRPs based on proportion of B_0 or its proxies in Canada (Marentette et al. unpublished working paper), this default policy guidance would be beneficial.

Table 2. Recommended proxies for B_{MSY}/B_0 and F_{MSY} based on $F_{X\%SPR}$ (see Section 2.2) based on productivity categories from MF (2011), and the intrinsic rate of population growth (r).

Productivity	r	B_{MSY}/B_0	$F_{MSY} \approx F_{X\%SPR}$
Very Low	-*	≥0.45	$F_{\geq 50\%}$
Low	<0.14	0.40	$F_{45\%}$
Medium	0.14-0.35	0.35	$F_{40\%}$
High	>0.35	0.25	$F_{30\%}$

* Defined based on M < 0.1 and age-at-maturation > 15 years.

2.2. THEORETICAL PROXIES FOR B_{MSY}

Per-recruit (also called "dynamic pool") models (Beverton and Holt 1957) are estimates of the lifetime expectation of the contributions of a single recruit to various metrics such as yield, biomass (or SSB), or egg production, and can be used as simple proxies for MSY reference points (DFO 2016). The advantage of using dynamic pool models is that they require less data because they do not consider recruitment dynamics and do not require a time series of catch or abundance data. They do, however, require assumptions to be made about growth of individuals and M (Kronlund et al. 2018) and are sensitive to changes in fishery selectivity (Wakeford et al. 2018). The disadvantage of per-recruit reference points is that they do not account for reductions in the number of recruits as SSB declines and therefore do not provide an internal basis for defining a threshold for recruitment overfishing (Sainsbury 2008). This has led to widespread use of 'rules of thumb' concerning what proportion of per-recruit reference points to use as a proxy for B_{MSY} (MF 2011). Per-recruit reference points are generally calculated using SSB as a measure of reproductive potential (e.g., SPR described below). However, SSB should be considered interchangeable with spawners, egg production, larval production or any other measure of reproductive potential (Brooks 2013).

The spawning potential ratio (SPR; Gabriel et al. 1989; Clark 1991; Mace and Sissenwine 1993) is defined as the SSB-per-recruit (φ) at a given constant, long-term F divided by the φ at long-term F = 0 (φ_0). $F_{X\%SPR}$ represents the F that results in X% of φ_0. The value of X that results in F_{MSY} is an inverse function of h (Williams 2002; Punt et al. 2008; Forrest et al., 2010). However, while F_{MSY} is a function of both h and M, once X is selected, $F_{X\%SPR}$ is estimated independent of h, although still sensitive to M (Forrest et al. 2018). $F_{40\%SPR}$ is commonly used as a proxy for F_{MSY} (Mace 1994; Clark 2002; Sainsbury 2008). F-based reference points can be interpreted in terms of biomass-based reference points as the equilibrium biomass resulting from long-term fishing at the constant specified F. Thus, the equilibrium biomass at $F_{40\%SPR}$ is a proxy for B_{MSY}; however, SPR values in the range of 30-50% are relatively common proxies for F_{MSY} (MF 2011; PFMC 2014). In New Zealand's operational guidance, $F_{40\%SPR}$ corresponds to an F_{MSY} for "medium productivity" stocks (Table 2), while the SPR for "low" and "high" productivity stocks are 45% and 30%, respectively (MF 2011). $F_{50\%SPR}$ has further been suggested by Sainsbury (2008) as a default proxy for F_{MSY} when h is unknown, since it provides a high fraction of MSY for h greater than 0.3. Dorn (2002) recommended fishing mortality rates in the range of $F_{40\%SPR}$ to $F_{60\%SPR}$ should be considered for data-limited West Coast rockfishes.

Additional theoretical proxies for B_{MSY} can be found in Appendix B.

2.3. LRPS BASED ON IMPAIRED RECRUITMENT

When a SRR is assumed, MSY reference points can be estimated (Section A.1); however, direct estimates of thresholds to avoid recruitment overfishing can also be made. These consist of non-parametric estimates based on the replacement F (Sissenwine and Shepherd 1987) or methods of estimating the biomass below which average recruitment declines or stock dynamics are highly uncertain (ICES 2018). For example, ICES guidance includes different methods for estimating LRPs based on the characteristics of the SRR. ICES (2021) defined six stock types based on SRRs and prescribed methods to estimate B_{lim} (equivalent to an LRP) for each type based on change points in segmented regressions or biomass estimates based on lowest observed biomass (ICES 2021a, Section A.1).

Non-parametric replacement F-based reference points can be used as a threshold for recruitment overfishing in the absence of a well-fitting SRR (Sissenwine and Shephard 1987), based on the theory that the persistence of a population requires that each recruited year class replaces the biomass of its parents (i.e., SSB) on average. The slope of a straight line through each point on the SRR plot and the origin (slope of $1/\varphi_F$) corresponds to an F that would be applied over the lifetime to those recruits in order to obtain SSB. The authors define recruitment overfishing as an F that reduces the SBB produced by a year class below the SSB of its parents on average. The slope of the replacement line is defined as the median ratio of recruitment to SSB. The mean long-term equilibrium SSB that results from fishing at F_{rep} (i.e., B_{rep}) is estimated as the SSB where the median replacement line intersects the SRR curve. In the absence of a SRR, the curve can be represented as the median recruitment (e.g., DFO 2002). Fishing at F_{rep} will result in B_{rep} on average, a proxy for B_{MSY}. A lower threshold for recruitment overfishing in terms of biomass can be defined as the equilibrium biomass from fishing at F_{rep90} which is the 90th percentile of the ratio of recruitment to SSB (Serebryakov 1991, DFO 2002). The F_{rep} and F_{rep90} reference points are consistent with the F_{med} and F_{low} defined by ICES (1988).

The hockey stick SRR (or segmented regression) allows an estimate of a threshold for recruitment overfishing by identifying a biomass below which the recruitment rate begins to decline (e.g., ICES 2021, Figure A.1). Various forms of the hockey stick SRR have been applied (e.g., segmented regression, ICES 2018; smooth hockey stick, Mesnil and Rochet 2010) and various methods of estimating the breakpoint have been used (e.g., grid-search method, Barrowman and Myers 2000). The shape of the hockey stick SRR is convenient because a biomass can be identified below which recruitment declines linearly, consistent with the definition of recruitment overfishing where there is no longer a compensatory buffer in juvenile survival (Hilborn and Walters 1992).

The biomass producing 50% of maximum recruitment (Myers et al. 1994) or 50% of the maximum recruitment predicted from the SRR (R_{max}) can also be estimated as a candidate LRP (DFO 2002; NAFO 2004). The choice of 50% is somewhat arbitrary and does not consider the variability in the SRR or *h* but has been used as a general rule of thumb for a threshold to recruitment overfishing. However, if mean recruitment is nearly constant across all levels of biomass (i.e., *h* is high) then 50% of maximum recruitment can occur at a low level of depletion (Myers et al. 1994) that could be associated with greater risk of serious harm.

2.4. LRPS BASED ON UNKNOWN STOCK DYNAMICS AT LOW BIOMASS

Uncertainty concerning population dynamics can be high at low stock sizes and in these cases, thresholds can be defined to delineate stock states with unknown dynamics (NAFO 2004, DFO 2006). The lowest biomass observed (B_{loss}) and from which there has been a recovery (often specifically designated as $B_{recover}$ or B_{rec}) are historical reference points that can represent this concept (DFO 2016, ICES 2021a, Rivard and Rice 2003). They are also easy to understand and communicate (DFO 2004) and can also be applied to empirical indicators (e.g., survey indices and CPUE time series). This reference point may be recommended for stocks with occasional large year classes (e.g., spasmodic stocks, ICES 2021a) or that have shown no evidence of impaired recruitment (non-informative SRR). These approaches are not suitable for stocks where recruitment appears to increase with stock size, nor stocks with narrow ranges of estimated biomasses (ICES 2021a). If used, the reference point should not be taken from recent years if the stock is declining (ICES 2021a).

The lowest biomass where “large recruitment” has been observed has also been used as an LRP (e.g., for spasmodic stocks, ICES 2021a). Large recruitment can be interpreted differently (e.g., the 50^{th} or 80^{th} percentile in a time series, van Deurs et al. 2021) and the determination of large recruitment may be determined by expert judgment (ICES 2015). When using percentiles of data distributions, to estimate reference points, van Deurs et al. (2021) showed that the length of the time series can strongly influence the reference point estimates.

A common assumption in using historical reference points, like empirical reference points, is that of stationarity (Kronlund et al. 2018, DFO 2004) and periods associated with B_{loss} or a “large recruitment” may correspond to favourable environmental conditions and/or low risk of excessive exploitation (DFO 2008). It may not be possible to determine whether stocks depleted to historic lows will again be able to rebuild from that level if the stock is expected to be under unfavourable productivity conditions (Rivard and Rice 2003). Alternatives should be considered if there is strong evidence of such conditions (DFO 2016) when knowledge of stock productivity allows (DFO 2008).

2.5. LRPS BASED ON RECOVERY FROM PERTURBATION

Resilience, or the ability to recover from perturbation, is arguably the stock attribute most pertinent to the risk of collapse but can be hard to define and measure (MF 2011). Lack of resilience could be defined as an impaired ability to exceed replacement and grow in the sense of stock rebuilding or stock recovery, e.g., to target levels capable of producing MSY (DFO 2016, Sainsbury 2008). Evaluation of impaired ability requires a time frame often linked to generation time, i.e., within one (NAFO 2004) or two (Rivard and Rice 2003) generation times under good conditions, or on timescales meaningful for human intergenerational equity (Sainsbury 2008), where generation time can be defined as the mean age of individuals contributing to the reproduction of the stock. An example of such an LRP would be S_{gen}, the spawner abundance that will result in recruitment equal to spawners at S_{MSY} (spawner abundance at MSY) in one generation in the absence of fishing under equilibrium conditions (DFO 2016). Care must be taken in defining LRPs based on recovery, as the estimate will depend on how the recovered state and time period are defined, as well as assumptions around productivity. Such LRPs can, however, be useful as a basis with which to compare the plausibility of other LRPs (Rice and Rivard 2003) where plausibility refers to whether estimates, assumptions or hypotheses are consistent with empirical data and ecosystem and population dynamics theory (e.g., past trajectories of stock indicators or productivity in light of fishing pressure and past or future environmental conditions).

2.6. SPATIAL REFERENCE POINTS

Spatial reference points may be most important for species with limited dispersal of early life-history stages (NAFO 2004). A threshold for serious harm was defined by a NAFO working group as a decrease in area of distribution of 75% (NAFO 2004), which is based on the assumption of a MacCall (1990) basin-model-like distribution pattern. This basin model assumption is that a reduction in spatial distribution will overestimate the reduction in population size. This is because density in the preferred habitat is hypothesized to increase with decreasing population size (i.e., the population will aggregate more densely in the preferred habitat or “the deepest part of the basin” in MacCall’s basin model distribution pattern). The 25% of the area of distribution would therefore maintain at least 25% of the population size.

Reuchlin-Hugenholtz et al. (2016) developed a methodology to estimate the threshold below which SSB declines increasingly quickly relative to a spatial indicator. They defined $HDA_{X\%}$ as the proportion of total survey tows in a given year that fall within the highest X% of non-zero tows over the entire time series distribution of tows. $HDA_{X\%}$ is therefore an index for areas containing the highest concentration of biomass, where values of X in the range of 2.5 to 15 were explored. Reuchlin-Hugenholtz et al. (2016) identified a concave relationship between standardized SSB and the standardized HDA indicator and found that across six groundfish stocks 0.2-0.3 HDA, a loss of 70-80% of high-density areas, consistent with the threshold defined above by NAFO (2004), was associated with potentially large declines in SSB.

2.7. TRAFFIC LIGHT APPROACH

The term “traffic light approach” (Caddy 2002), has been used to describe a management framework that uses a system of green, amber, and red lights (e.g., Halliday et al. 2001) or simply red and green lights (e.g., ICES 2018) to categorize multiple indicators of the state of the fishery. A threshold for serious harm (or an undesirable state) is determined for each indicator. In the absence of an objective basis for defining a threshold for an indicator, expert judgement can be used (Halliday et al. 2001; Harford et al. 2021). Some form of integration of indicators is required to support an estimate of overall stock status. The integration can involve scaling the indicators to make them comparable and using some operation to combine the indicators (Halliday et al. 2001).

Advantages of the traffic light approach include its simplicity and ability to accommodate multiple indicators including those that are directly observable. The traffic light approach may be considered in situations where multiple data sets cannot be combined statistically but can inform on different aspects of stock status that are relevant for management objectives (Harford et al. 2021). Multiple indicators can be appealing when a single indicator could be ambiguous (e.g., a decrease in mean length could be driven by a strong recruitment event or by a loss of older individuals in the population). The challenges may be in defining a method of integrating the indicators into an overall metric of stock status. Various approaches have been used, including using the proportion of indicators in the red category and using a weighting method to combine multiple indicators that are proxies for the same variable (Harford et al. 2021).

2.8. LRPS FOR SEMELPAROUS STOCKS

For semelparous species, SRRs can be used to describe the entire life history of a stock. Estimates of the spawner abundance at MSY (S_{MSY}), the optimal harvest rate at MSY, and the spawner abundance that produces maximum recruitment (R_{max}) can be directly estimated from the SRR parameter estimates (Hilborn and Walters 1992; Table 3). Candidate LRPs can be defined based on a proportion of S_{MSY} (e.g., 0.4 S_{MSY}), the spawner abundance at 50% R_{max}, or S_{gen}, the spawner abundance that results in recruitment equal to spawners at MSY in one generation in the absence of fishing at equilibrium conditions (e.g., Holt et al. 2009; Chaput 2015). Reference points based on S_{MSY} and S_{gen} have been adopted under Canada's Wild Salmon Policy (WSP) (Holt et al. 2009) and have also been used as operational control points (OCPs) for salmon, below which harvest is limited to low levels (DFO 2016).

Table 3. Reference points for semelparous stocks estimated from the Beverton-Holt (BH) and Ricker stock recruitment relationships parameterized as in Equations A.1 and A.2 in Appendix A. (Adapted from Hilborn and Walters 1992, based on the Equation A.1 and A.2 parameterizations).

Reference Point	BH (Equation A.1)	Ricker (Equation A.2)*
S_{MSY}	$\frac{\sqrt{\alpha}-1}{\beta}$	$\frac{\ln(\alpha)}{\beta}\left(0.5-0.07\ln(\alpha)\right)$
R_{max}	$\frac{\alpha}{\beta}$	$\frac{\alpha}{\beta}e^{-1}$
u_{MSY}	$1-\sqrt{1/\alpha}$	$0.5\ln(\alpha)-0.07\left(\ln(\alpha)\right)^2$

***** S_{MSY} and u_{MSY} formulas are approximations that are valid over $0 < \alpha < 3$ (Hilborn 1985)

2.9. LRPS FOR PACIFIC SALMON

Pacific salmon (*Oncorhynchus* spp.) are unique among marine fish species due to their high levels of intraspecific diversity that gives rise to a large range in data availability, considerations, and approaches for assessments and LRP development. For Pacific salmon, LRPs have been developed to align with Canada's Wild Salmon Policy (WSP) objective of preserving adaptive diversity of salmon at the scale of conservation units (CUs), which are nested within major stocks, or salmon "Stock Management Units", SMUs (Holt et al. 2023a, Holt et al. 2023b). Methods for evaluating status of CUs are well established using biological benchmarks to delineate three zones of status: green, amber, and red representing increasing conservation concern and management intervention. Given diversity in data availability and types among CUs, a traffic light approach (Section 2.7) is used to estimate status where statuses on multiple metrics are combined into an overall assessment of green, amber, or red (Holt et al. 2009; DFO 2024).

For the purposes of the FSP, LRPs can be identified from the proportion of CUs that have status above the red zone for WSP status assessments (Holt et al. 2023a). This approach provides some consistency with status assessments already produced under the WSP, and can inform management decisions for harvest, habitat, and hatcheries that often occur at finer, CU scales. To supplement the default approach, LRPs based on metrics of aggregate abundances for the entire SMU have also been proposed (Holt et al. 2023a; Holt et al. 2023b), which may be required for fisheries management purposes in some cases. For this approach, LRPs are derived to have a desired probability of all component CUs being above red status and assume a relationship between aggregate abundance and the proportion of CUs that have status above the red zone.

These LRP methods may be adaptable to other species and assessment management contexts where data are limited, traditional estimation approaches cannot be applied, or biological units of assessment are nested within major stocks. The “traffic light” approach is particularly useful for species where theoretical or historical abundance-based metrics are not estimable or applicable, or where data are limited (Dowling et al. 2015).

2.10. LRPS FOR ATLANTIC SALMON

LRPs for Atlantic salmon (*Salmo salar*) in Canada have been defined on the basis of conservation of the salmon population by setting escapement goals (DFO 2015). Candidate LRPs consistent with the PA Policy have been defined based on the SRR in terms of S_{MSY}, S_{gen}, the spawner abundance at 50% R_{max}, or spawning abundance that results in less than 25% probability of recruitment being less than 50% of R_{max}. (DFO 2015). In Atlantic salmon populations there is a strong sex bias in the sea age-at-maturity (O’Connell et al. 2006) and differences in the abundances of sea age groups in annual returns to a river (DFO 2015). Egg deposition rates are therefore commonly used as the indicators of stock status for Atlantic salmon. The LRP most recently defined for Atlantic salmon in Canada was the egg deposition rate (i.e., number of eggs per area of habitat) that results in less than a 25% chance that the realized smolt production from freshwater would be less than 50% of R_{max} (DFO 2015; 2018b).

2.11. LENGTH-BASED APPROACHES

The length-based approaches include length-based indicators of stock status and length-only assessment methods that generally only allow estimation of *F*-based reference points (Pons et al. 2020). Some length-based approaches do, however, include estimation of biomass-based reference points (e.g., Froese et al. 2018) or alternative indicators of stock status (e.g., based on SPR). The theory behind length-based indicators is that there is a decline in average length in the catch because of high fishing mortality that truncates the size distribution, or size-selective fishing mortality where larger individuals are removed from the population. Most length-based approaches assume equilibrium conditions and the methods are influenced by strong year classes (reducing mean length and suggesting high mortality) so it is critical to understand whether extreme year-classes are influencing the length distributions (ICES 2018).

2.11.1. Length-based Indicators

Froese (2004) first introduced three length-based indicators based on the length frequency distribution of the catch with targets:

- P_{mat}, the percent of mature fish in the catch (target 100%),
- P_{opt}, the percent of fish with optimum length (size at which the highest yield from a cohort occurs) in the catch (target 100%), and
- P_{mega}, the percent of “mega-spawners” (large mature fish) in the catch (target 0%), 30-40% mega-spawners represents a healthy age structure and < 20% is a concern.

These metrics were intended to avoid growth and recruitment overfishing; however, the metrics have no quantitative linkage to metrics of stock status (e.g., B_{MSY} or B_0). Cope and Punt (2009) explored the relationships between these indicators and measures of *F* and SSB using simulations from an age-structured model. They found that each metric showed a wide range of possible values depending on selectivity, *h*, and the ratio of length-at-maturity to optimal fishing length and that a single metric was not sufficient to ensure sustainable fishing. Even with P_{opt} = 100%, major decreases in biomass can still occur depending on the selectivity. Cope and Punt (2009) therefore provided a decision tree for interpreting stock status when the fishery is not operating on the optimal selectivity. The authors proposed a new metric ($P_{obj} = P_{mat} + P_{opt} + P_{mega}$) to identify reference points in terms of biomass relative to a target biomass, based on selectivity type which is determined by the value of P_{obj}.

A set of length-based indicators (LBIs) is used by ICES to assess data-limited stocks, which follows the traffic light approach (Caddy 2002; Section 2.7). The traffic light approach used by ICES for LBIs is a system of red and green lights to categorize multiple indicators in relation to desired stock status states. The ICES LBIs require the following data: length-at-maturity (L_{mat}); asymptotic length (L_{inf}); catch-at-length by year; and the weight-length parameters (*a* and *b*) from the catch. The LBIs are grouped in terms of indicators of conservation/sustainability, optimal yield, and MSY (Table 4). The approach assumes equilibrium conditions and that the fishery selectivity ogive is flat-topped (logistic) (ICES 2018). The optimal length (L_{opt}) is estimated assuming M/k = 1.5, where k is the von Bertalanffy growth rate parameter. However, if direct estimates of M and k are available then reference points L_{opt} and $L_{F=M}$ can be estimated as:

$$L_{opt} = L_{inf}\frac{3}{3+M/k} \tag{1}$$

$$L_{F=\gamma M.k=\theta M} = \frac{\theta L_{inf} + L_c(\gamma+1)}{\theta+\gamma+1} \tag{2}$$

where the parameters are as defined in Table 4 and γ and θ are user-defined proportions of M used as proxies for F_{MSY} and k, respectively, when assumptions of $F_{MSY} = M$ and M/k = 1.5 are not supported. The expected values of the LBIs for optimal yield and MSY in theory reflect length-frequency distributions of the catch under equilibrium conditions fishing at F_{MSY}. The indicators are sensitive to the assumed values of L_{mat} and L_{inf}, and sensitivity analyses are recommended to evaluate the influence of these values on status estimates (Shephard et al. 2018). The suitability of the assumed ratio of M/k = 1.5 should also be evaluated. Although the ratio of M/K is believed to be less variable across stocks/species than M (Prince et al. 2015), the ratio can be outside the bounds of 1.2 to 1.8 for many exploited species (Hordyk et al. 2019). Length-frequency distributions should be inspected for evidence of a unimodal distribution to support the assumption of equilibrium conditions (ICES 2018).

Candidate reference points for each indicator are provided in Table 4. An overall estimate of stock status would be obtained using a traffic light approach (Section 2.7) or a weight-of-evidence approach (Section 2.12).

Table 4. Length-based Indicators used by ICES (2018).

Indicator Type	Indicator	Calculation	Reference Point*	Expected Value of the Ratio: Indicator/ Reference Point
Conservation (large individuals)	$L_{max5\%}$	Mean length of largest 5%	L_{inf}	>0.8
	$L_{95\%}$	95th Percentile	L_{inf}	>0.8
	P_{mega}	Proportion of individuals > 1.1 L_{opt}	0.3 – 0.4	P_{mega} > 0.3

Indicator Type	Indicator	Calculation	Reference Point*	Expected Value of the Ratio: Indicator/ Reference Point
Conservation (juveniles)	$L_{25\%}$	25th Percentile	L_{mat}	>1
	L_c	Length at 50% of modal abundance	L_{mat}	>1
Optimal Yield	L_{mean}	Mean length > L_c	L_{opt} = 2/3 L_{inf}	≈1
	L_{maxy}	Length class with maximum biomass in catch	L_{opt} = 2/3 L_{inf}	≈1
MSY	L_{mean}	Mean length > L_c	$L_{F=M}$ = 0.75 L_c + 0.25 L_{inf}	≥1

*L_{opt} and $L_{F=M}$ reported for M/k = 1.5

2.11.2. Length-based Methods for SPR Ratio and Biomass-based Indicators

A length-based SPR (LBSPR) method was developed by Hordyk et al. (2015a, 2015b) and has been adopted by ICES (2018) in their guidance for data-limited stocks. For this approach, the indicator of stock status is the annual estimated %SPR and an estimate of F/M (a proxy for F/F_{MSY}) is also generated. The SPR is estimated from a representative sample of the size distribution from the catch and some assumptions about the life-history characteristics of the species. The LBSPR method assumes equilibrium conditions and that the length distribution data are a representative sample from the exploitable population. The method uses the ratio of M/k (less variable across stocks/species than M; Prince et al. 2015), estimates of asymptotic length (L_{inf}) and the von Bertalanffy growth parameter k, maturity-at-length (i.e., length at 50% and 95% maturity), the weight-length relationship parameters (a, b), and fishery selectivity-at-length (assumed to be knife-edged, logistic, or domed shaped). SPR is estimated using maximum likelihood estimation to find values of F/M and selectivity-at-length that minimize the difference between observed and expected length composition of the catch. The annual SPR can be estimated and used as an indicator of stock status. A 40% SPR is a common target reference point (i.e., $F_{40\%}$ is a proxy for F_{MSY}; see Section 2.2) although the percentage depends strongly on the assumed resilience of the stock (Table 1). Goodyear (1993) provides a range of 20-30% SPR as minimum acceptable levels above which stocks maintain acceptable productivity, the minimum value of this range consistent with 0.5 of a target of 40% SPR. Mace and Sissenwine (1993) similarly provide a range of 20-35% SPR as a threshold for recruitment overfishing and recommend 30% SPR as an LRP when stock resilience is unknown.

A length-based integrated mixed effects method (LIME; Rudd and Thorson 2018) also uses length composition data from the catch and life-history data to estimate F and SPR but does not have the equilibrium assumption of LBSPR. LIME uses mixed effects to estimate changes in recruitment and F separately over time and estimates one selectivity curve for the entire time series (Rudd and Thorson 2018). All other data requirements are the same as LBSPR.

A length-based Bayesian biomass estimation method (LBB; Froese et al. 2018) can estimate B/B_0 and B/B_{MSY} from the analysis of length frequency distributions of the catch. The LBB method relies on the von Bertalanffy growth equation. The LBB framework is based on relative rates (instead of absolute rates) which reduces the number of parameters. An increase in length of a cohort is used as a proxy for time and, by using ratios instead of absolute values, the units of time and biomass cancel out (Froese et al. 2018). The rates M/k and F/k are used in the LBB method to estimate F/M (here M is a proxy for F_{MSY}) and B/B_0. Hordyk et al. (2019) note that the LBB method is highly sensitive to the assumed asymptotic length.

Pons et al. (2020) simulation tested the LBSPR, LIME, and LBB methods for three life-history scenarios (short-lived, medium-lived, and long-lived) crossed with three harvest trends (F increasing to a maximum then declining, F increasing then remaining constant, and F constantly increasing), and three depletion scenarios (final depletion 0.2, 0.4, and 0.6) to evaluate bias and precision. In general, LIME was the least biased method. LBSPR generally underestimated harvest rates. For short-lived species, LIME performed best, whereas LBB had the most variability in F estimates. For medium-lived species, LIME performed best and results for LBSPR and LBB were similar, but LBSPR was more precise. Pons et al. (2020) showed that LBB was highly biased and imprecise for long-lived species. This result was also noted by the authors of LBB, who didn’t recommend using it for long-lived species (Froese et al. 2018). LIME was also highly imprecise for long-lived species, a result also noted by the authors of the method (Rudd and Thorson 2018). LIME was, however, the least biased among the three methods for all life-histories. The general recommendation of Pons et al. (2020) was to use LIME except for long-lived species, for which LBSPR was recommended. However, a long time series of data is generally needed to draw conclusions for long-lived species using this method.

2.12. STATUS BASED ON EXPERT JUDGEMENT

The use of expert judgement alone to assign stock status has sometimes been invoked where all other methods appear to be infeasible. Expert judgement has been used in data-rich scenarios, where expertise is widely applied in Bayesian frameworks to develop stock status priors (Chrysafi et al. 2019), where expertise is used to select a minimum biomass that had “good” recruitment (ICES 2015), to select an undesirable historical biomass to be avoided in future (e.g., DFO 2019b), or to integrate estimates of status over multiple, often conflicting metrics, in a “composite” approach (e.g., Grant and Pestal 2012). Methods of expert elicitation vary from formalized Delphi techniques (Rowe and Wright 2001), where experts are anonymously and iteratively consulted to provide judgements on a given topic, interspersed with discussion of expert inputs at the group level; to less intensive but standardized software tools displaying available information (Chrysafi et al. 2019). Regardless of method, rigour should be applied in the collection of expertise to ensure expert judgement is both transparent and repeatable (Drescher et al. 2013).

Another example of expert judgement methodologies is the use of weight-of-evidence frameworks. Weight-of-evidence is used in risk assessment in a variety of fields and can be summarized into general principles as follows, with each activity supported by sufficient documentation (Health Canada 2018):

- gathering all available evidence,
- assessing individual studies for inclusion/exclusion based on screening criteria (relevance, quality, reliability, etc.),
- assembling lines of evidence (i.e., groups of related studies of data sources),
- assessing lines of evidence for their strength or robustness in support of, or against, a particular hypothesis, and
- integrating multiple lines of evidence to an overall conclusion by either qualitative or quantitative processes, including the use of weighting schemes.

In 2009, Australia's Department of Agriculture, Fisheries and Forestry launched the Reducing Uncertainty in Stock Status (RUSS) project, with two streams. Stream 1 developed a weight-of-evidence framework to systematically classify stock status of data-limited stocks (Larcombe et al. 2015), which was subsequently used to inform stock status designations for the Status of Australia's Fish Stocks (SAFS) report (Flood et al. 2016). Stream 2 conducted closed-loop simulations to evaluate whether data-limited harvest strategies could still achieve the intent of applicable policies (Haddon 2012). Stream 2 was motivated by the greater array of input and output controls typically used for data-poor stocks, including indicators that differ from "traditional" strategies relying on estimates of fishing mortality and stock size to set catch limits.

The RUSS weight-of-evidence framework aimed to provide structure for a scientific review and interpretation of indicators of biomass and *F* (evaluated separately), arriving at a status determination for each attribute through cumulative weight-of-evidence, in a transparent and reproducible way (Larcombe et al. 2015). It involved:

- compiling evidence of:
    - stock and fishery attributes related to productivity, *M*, distribution, aggregation, mobility, structure, whether there is a targeted fishery, how many fisheries, and whether it is a multi-species fishery,
    - empirical indicators such as trends in catch, effort, catch rates, spatial distribution of the fishery and age- and size-structure of the catch,
    - fishery independent surveys, if available,
    - stock assessment models, if available, and
    - risk assessments, such as the sustainability assessment for fishing effects (SAFE) approach or productivity-sensitivity analyses (PSA; but see Hordyk and Carruthers 2018).
- documenting and weighing evidence, by:
- identifying key lines of evidence, their interpretation, and the rationale for using them;
- describing whether different weights were assigned to different lines of evidence, and why;
- identifying inconsistent lines of evidence, including whether these were weighted differently; and
- identifying key areas of uncertainty;
- concluding by assigning status, including a peer review component.

Regardless of framework, it is important to note that the experience level of experts and the "true" stock status can both impact the performance of expert judgement-based methods of stock status assignment. Expert judgement has been shown to perform best for stocks with intermediate status (around B_{MSY}), whereas experts tended to under-estimate stock status for underexploited stocks and over-estimate status of depleted stocks. Furthermore, less experienced experts with little to no background in stock assessment can be both more biased, and more certain of their choices (Chrysafi et al. 2019, Chrysafi and Cope 2019).

3. PROPERTIES OF LRPS

3.1. UNCERTAINTY

Uncertainty in the context of fisheries can be defined as "the incompleteness of knowledge about the state or processes (past, present, and future) of nature" (FAO 1995). Six types of uncertainty in fisheries management have been defined: uncertainties associated with process, observation, model, estimation, implementation, and institutions and these are reviewed in

(Francis and Shotton 1997). The first four types of uncertainty can be considered scientific uncertainty and the latter two can be considered management uncertainty (e.g., Privitera-Johnson and Punt 2020). The estimation uncertainty in LRPs and estimates of stock status relate to scientific uncertainty and result from the combined effects of process, observation, and model uncertainties. The process uncertainty results from the underlying stochasticity or natural variability in population dynamics such as annual variability in recruitment. Observation uncertainties arise from the process of data collection resulting in measurement or sampling error from observing a sample from a statistical population. Model uncertainty refers to assumptions about structural forms within the model (e.g., form of the SRR, shape of the selectivity ogive, assumptions about whether parameters are age- or time-varying). Some methods for calculating LRPs include estimation uncertainty and some do not. When reference points are estimated within an assessment model, estimation uncertainty can be accounted for in internally estimated reference points. In assessment models where key parameters that determine reference points are fixed (e.g., h and/or M), uncertainty in reference points will be underestimated and the reference points will potentially be biased (Mangel et al. 2013).

Estimates of LRPs can also be made outside of the assessment model (e.g., by inputting stock assessment outputs into reference point software packages), but assumptions like the form of the SRR may not be consistent with the original assessment model (ICES 2022). Uncertainty in LRP estimates outside of an assessment model can be estimated using various methods (e.g., Monte Carlo simulations, Braccini et al. 2015; bootstrapping, Cadigan 2013, ICES 2022; closed-loop simulation, Merino et al. 2020). However, these methods may not maintain the correlations among model parameters (Trijoulet et al. 2022). When model estimated stock-recruit pairs are treated as independent observations in subsequent analyses, spurious relationships and underestimates of uncertainty can introduce bias in reference point estimates (Link 1999; Maunder and Punt 2013; Brooks and Deroba 2015).

Characterization of the uncertainty in stock status is critical because the statistical properties of uncertainty captured in an estimate of stock status affect conclusions about stock status relative to reference points and the choice of management measures aimed at meeting objectives related to reference points. The PA Policy is focused on tolerance for stock decline. With respect to the LRP, the management objective is to avoid biomass levels below the LRP with high probability.

Sensitivity analyses can be used to qualitatively characterize uncertainty by demonstrating model behaviour and assess the robustness of model results to specific model formulations or alternative data sources (Privitera-Johnson and Punt 2020). Often, a base or reference model is selected based on expert judgement, and advice is based on this model. Sensitivity analyses can be used to evaluate the consequences of uncertainty from alternative model assumptions on model outputs and advice. Model uncertainty can also be captured in an estimate of stock status by combining estimates from multiple models using a model-averaging or ensemble approach (Section 4.2). These could be alternative assessment models or operating models that represent alternative hypotheses in a closed-loop simulation approach.

When uncertainty in stock status can be estimated, a probability distribution for current or projected biomass can be generated and the probability of being above the LRP can be estimated. There is an underlying probability distribution for current or projected biomass and also for the estimate of the LRP. To capture this variability, the ratio of biomass to the LRP is expected to have greater precision and accuracy than using the absolute values in terms of biomass (NAFO 2004).

3.2. RELIABILITY

Reliable estimation of LRPs or stock status metrics can be interpreted as being acceptably robust (considering consistency, variance, and bias) to key uncertainties and assumptions in the advice framework (Marentette et al. unpublished working paper). The true values of reference points are unknown so evaluations of the accuracy of estimates are conducted in simulation tests of systems with known values. Historical fishing pressure and the availability of data influence the reliability of reference point estimates. In a simulation study, Magnusson and Hilborn (2007) identified that a large decrease in abundance is the most important feature in the fishery time series for informing reliability of reference points, followed by a time series with contrast in harvest rates. Reliability of LRP estimates depends on the method used to estimate uncertainty in the estimates. An evaluation of methods to quantify uncertainty can also be conducted using simulation testing. For example, Magnusson et al. (2013) conducted a simulation study to evaluate bias and width of confidence intervals for reference points using three methods of quantifying reference points in an age-structured model. Simulation studies have also been used to estimate the bias and precision of data-limited methods to estimate metrics like B_{MSY} and SPR (e.g., Pons et al. 2020).

The reliability of estimates of B_{MSY} depends on how well h is estimated. Estimates of h in a simulation study were more reliable when age-composition data and an index of abundance data were available from time periods with very low abundance (Magnusson and Hilborn 2007) which resulted in estimates of recruitment at low biomass. Given the challenges in estimating h, priors for h are commonly used when fitting assessment models (e.g., Myers et al. 1999; Mangel et al. 2010). Estimates of h for similar stocks from meta-analyses (e.g., Thorson et al. 2020), or estimates from life-history characteristics (Wiff et al. 2018) can be used to guidance estimates or prior distributions for h. Sensitivity analyses can be conducted to evaluate the influence of model parameters on reference point estimates.

3.3. PLAUSIBILITY

The plausibility of candidate indicators of stock status and LRPs refers to whether estimates, assumptions or hypotheses are consistent with empirical data and ecosystem and population dynamics theory (e.g., past trajectories of stock indicators or productivity in light of fishing pressure and past or future environmental conditions). When comparing multiple candidate LRPs, an evaluation of the plausibility of each candidate approach may support LRP selection. For example, proportions of model based estimates of B_0 and B_{MSY} can be compared to $\mathrm{B_{loss}}$ and $\mathrm{B_{recover}}$ or with past states associated with low surplus production to assess whether recovered from these candidate values have occurred in the past.

4. STOCK STATUS METRICS

4.1. HIGH VARIABILITY

Approaches for addressing high variability in estimates of stock status include using smoothers (e.g., *loess*: locally-weighted scatterplot smoothing or a 3- or 5-year moving average) which are commonly used for empirical indices of stock status for Canadian stocks (Marentette et al. unpublished working paper)). Smoothing methods can also be used for model-based estimates of stock status. For example the R package for the LBSPR method uses a smoother to reduce annual variability in SPR and *F*/*M* (Hordyk 2021).

4.2. SINGLE STOCK STATUS FROM MULTIPLE MODELS

4.2.1. Ensemble Approaches

An ensemble approach (also known as model averaging) is any method of combining inference for quantities of interest across multiple models (e.g., Dietterich 2000; Anderson et al. 2017; Jardim et al. 2021). In stock assessment, those quantities of interest might be the LRP or stock status with respect to the LRP. An ensemble can combine models with, for example, alternative structural assumptions, distinct modeling platforms, parameters fixed at different values, or models with alternative priors. Any models that produce the same quantity of interest can be combined in an ensemble unless the ensemble relies on statistical weighting (e.g., based on Akaike information criterion [AIC], Bayes factors), in which case the same statistical framework must be used.

When using an ensemble of models, a decision must be made about how to assign weights to the component models. Options include:

- equal weighting;
- tactical weighting (e.g., based on expert opinion or historical performance);
- weighting based on model probabilities (e.g., Bayes factors);
- weighting based on information theoretic values (e.g., AIC); or
- weighting based on predictive ability.

An additional more complex form of weighting by a second-level model fit to known or trusted data is also possible ("superensembles", Anderson et al. 2017). Model selection can be thought of as an extreme version of ensemble modeling where one model is assigned all the weight (Jardim et al. 2021).

There are several issues to consider when choosing which models to include in an ensemble. First, it is critical that all component models are plausible and satisfy model assumptions (e.g., residual distributions, posterior predictive checks, convergence). Then, there are issues related to the composition of the models in the ensemble. Issues of model composition include too little overlap among models, too much overlap among models, or some combination of the two (e.g., "clumping" among models). With too little overlap, there is a risk of obscuring distinct plausible realities that would require alternative management actions—the best answer isn't necessarily in the middle (e.g., Anderson et al. 2017, Maunder et al. 2020). Solutions could include subgroup ensembles or not using ensembles at all. With too much overlap, the benefit of using an ensemble approach may be reduced, since the component models may not be representing a sufficient diversity of structural assumptions. This may or may not be an issue. However, if too much overlap occurs within groups of models, it can result in undesired unequal weighting among broader hypotheses. Solutions include refining the composition of the ensemble so that model groups are based on broader hypotheses or considering a hierarchical framework to group hypotheses and assign weights (Maunder et al. 2020).

If combining multiple models to derive a single estimate of the LRP or status with respect to the LRP, several best practices are recommended:

- Do not use ensembles as an excuse to avoid rigorous model validation of the component models.
- Make decisions about ensemble composition prior to fitting models to data if possible.

- Consider models that can expand or simplify based on the data if possible (e.g., estimating effects as random effects or letting parameters vary according to some constrained function) rather than including multiple models in an ensemble.

4.2.2. Closed-Loop Simulation

Closed-loop simulation is a process that involves evaluating the performance of alternative management procedures (MPs, algorithms for making management recommendations) through simulation to assess trade-offs among multiple fishery and conservation objectives (Butterworth and Punt 1999; Punt et al. 2016). This process is commonly called management strategy evaluation (MSE); however, we have reserved the term "MSE" in this paper to represent the full stakeholder-driven process. Closed-loop simulation is the technical modeling component of MSE, which can be used as part of a full consultative MSE process or on its own. In closed-loop simulation, multiple operating models (OMs) are typically used to capture a range of uncertainty in stock and fishery dynamics; however, a single OM can be used. Each OM has its own implicit biomass, reference points, and stock status under the assumptions and hypotheses of the OM. The objective of closed-loop simulation is to identify MPs with acceptable performance relative to management objectives, and not specifically to estimate the stock status relative to the LRP.

There has been a movement towards closed-loop simulation approaches in stock assessment and fisheries management worldwide (e.g., de La Mare 1998; Butterworth and Punt 1999; Plagányi et al. 2007; Rademeyer et al. 2007; Berkson and Thorson 2015; Geromont and Butterworth 2015; Carruthers et al. 2014; Punt et al. 2016). The closed-loop simulation approach can be used to provide management advice consistent with the PA Policy objectives of avoiding the LRP with a high probability and can be used as a tool to examine the consequences of alternative management measures aimed at sustainable management or rebuilding depleted stocks (DFO 2019a). Advantages of closed-loop simulation include the requirement for an explicit statement of conservation and fishery objectives, against which performance of alternative MPs are measured; and the ability to integrate testing of the performance of MPs across a range of alternative hypotheses about stock and fishery dynamics represented in the OMs. For this latter reason, closed-loop simulation is also a useful tool for evaluating performance of MPs in data-limited situations (Carruthers et al. 2014; Carruthers and Hordyk 2019; DFO 2020a; Anderson et al. 2021; Haggarty et al. 2022). Simulation testing of MPs allows for evaluation of performance across a range of uncertainties in fish stock and fishery dynamics, observation error, estimation error, and implementation error (e.g., Kell et al. 2006; Carruthers et al. 2016). While OMs are best conditioned on available data, a strength of closed-loop simulation in the data-limited context is that the OMs can also explore the robustness of MPs to plausible stock and fishery dynamics based on expert knowledge (Carruthers and Hordyk 2019; Anderson et al. 2021).

Reference points and stock status are an integral component of closed-loop simulation; however, the concept of an LRP may serve multiple roles or definitions in closed-loop simulation:

1. A component of a measurable management objective (e.g., greater than 95% probability of SSB > LRP over the long-term) that is defined in each OM and which is used to evaluate performance of MPs relative to the objective.
2. As an explicit metric of stock status to be used for determining whether a rebuilding plan is triggered under s 6.2 of the FSP.
3. An operational control point (OCP) in an MP (e.g., as used in the provisional HCR in DFO's PA Policy; DFO 2009).

Each of these three roles is explored below.

4.2.2.1. LRP Role 1: Component of Management Objectives and Performance Metrics

The first role of the LRP is as a fully specified objective in the closed-loop simulation framework, which includes a probability and a time frame. For example, the operational objective may be "P(SSB > LRP) > 95% in each and every year of the long-term component of the projection period." The LRP in this context is defined as a theoretical threshold consistent with the objective to avoid serious harm to the stock (e.g., as a fixed proportion of B_{MSY} or B_0, or as a minimum historical biomass).

Best practice involves separating OMs into reference and robustness sets of OMs (Rademeyer et al. 2007; Punt et al. 2016; Anderson et al. 2021), where the reference set of OMs consists of the most plausible hypotheses and the robustness set of OMs consists of hypotheses that are less likely but cannot be eliminated entirely and may have high consequence. The specific values of LRPs are unique to each OM and simulation replicate under consideration in the framework. In this role, the LRP is therefore an implicit component of the performance metrics used to evaluate the MPs within each OM and stock status is not typically explicitly reported in the advice. The advice provided is with reference to which MPs meet management objectives and the trade-offs that ensue relative to the other objectives.

In closed-loop simulations with multiple OMs, OMs may be weighted, where weights can be based on qualitative criteria (Butterworth et al. 1996), or may be quantitative based on model-selection criteria (e.g., AIC). However, Punt et al. (2016) urged caution in using model selection criteria to weight OMs unless there is confidence in the reliability of the likelihood function, which is unlikely to be the case in data-limited situations. Placing less plausible OMs into the robustness set may be one alternative to weighting of OMs. For example, Anderson et al. (2021) calculated performance metrics on average across an unweighted set of OMs in the reference set and presented performance metrics from each OM in the robustness set calculating performance metrics across operating models (OMs) separately. ICES (2020) identify two general approaches for calculating performance metrics across OMs:

- International Whaling Commission (IWC) approach, and
- Commission for the Conservation of Southern Bluefin Tuna (CCSBT) approach.

The IWC approach considers a manageable number of OMs and presents probabilities separately for each OM. MPs must meet mandatory management objectives related to the LRP for each OM in the reference set (ICES 2020). This has been termed "satisficing" (Miller and Shelton 2010), which simplifies the results set by screening out MPs that do not meet predefined minimum performance criteria (see also Anderson et al. 2021). The CCSBT approach can be applied with a greater number of OMs (but also can be applied to a small number of OMs) in a reference set and involves applying a weighting to each OM (e.g., based on fit likelihood or expert judgement). Performance statistics using the CCSBY approach are calculated across all weighted OMs in the reference set (ICES 2020).

There are also multiple ways to quantify the probability of dropping below the LRP in a performance metric. ICES (2013) describe three approaches that have been used in the past:

- average annual P(SSB < LRP) over the projection period,
- P(SSB < LRP) at least once during the projection period, or
- maximum annual P(SSB < LRP) over the projection period.

These probabilities can be calculated across all OMs (e.g., CCSBT approach) or separately for each OM (e.g., IWC approach) to evaluate performance. The choice of approach will affect the values of performance metrics, which in turn will affect MP selection criteria such as trade-offs with other performance metrics.

4.2.2.2. LRP Role 2: Metric of Stock Status

Given that closed-loop simulation focuses on LRP Role 1, a challenge under the FSP is how to estimate stock status and determine whether a rebuilding plan should be triggered. The FSP require the reporting of a single estimate of stock status for each candidate or prescribed fish stock. It is sometimes proposed to use an estimated LRP and stock status from the conditioned OMs using any of the methods discussed in this document. When a closed-loop simulation framework uses a single OM, this may be appropriate; however, when multiple OMs are used, these OMs may not represent a balanced set of models to combine or have a single best model to choose. In data-limited cases, a closed-loop simulation approach may have been chosen because estimated status using alternative methods was considered unreliable (e.g., Carruthers and Hordyk 2019, Anderson et al. 2021). In these cases, it is possible an MP can be selected that meets conservation and fisheries objectives with a high degree of confidence even if status is not well estimated and using the OMs to report status would undermine this form of science advice.

Multiple approaches to defining an indicator of stock status can be used when a closed-loop simulation framework is being used to provide harvest advice. A single (e.g., base case) OM can be selected to inform stock status that may have the “best” estimates of model parameters (e.g., Pacific herring, *Clupea pallasii,* DFO 2021); however, the full range of uncertainty in status will not be captured by a single OM. When OMs have been weighted based on model plausibility (e.g., CCSBT approach of calculating performance across OMs), or can be assumed equally plausible, an ensemble approach using the weighted average stock status among OMs can be used (e.g., Sablefish, *Anoplopoma fibria*; DFO 2020b). Although an estimate of stock status can be obtained from the OMs over the historical time period using this approach, updates of stock status would not be estimated during the period for which an MP is selected and applied. A trigger for changes in stock status during this period could be identified through exceptional circumstances in the framework process (e.g., Carruthers and Hordyk 2019; DFO 2018a), which could result in an evaluation of the MP and reconditioning of OMs.

When OMs may not represent a balanced set of models to combine or a single best model cannot be determined (e.g., data-limited scenarios, or the IWC approach of calculating performance across OMs), stock status may be determined using a weight-of-evidence approach (Section 2.12) based on the conditioned OMs (e.g., Inside and Outside Yelloweye Rockfish, *Sebastes ruberrimus*, DFO 2020a, DFO 2020c, Anderson et al. 2021). This follows the advice of Kronlund et al. (2021), who drew on precedent by Health Canada (2018), for a weight-of-evidence approach. If the conditioned OMs suggest a high probability of a fish stock being below its LRP across a range of plausible OM assumptions or if the best performing MPs have a low probability of maintaining the stock above the internal LRP (Role 1) in the near future, this contributes to the weight-of-evidence that can trigger a rebuilding plan.

An alternative approach to reporting stock status when a closed-loop simulation framework is used is to use an empirical indicator (e.g., survey index or empirical index of biomass), outside of the analytical modelling framework. This approach has been used for Western-component Pollock (*Pollachius virens*; DFO 2011), Atlantic herring (*Clupea harengus*; DFO 2022b), and Atlantic halibut (*Hippoglossus hippoglossus*; DFO 2023b) where a survey index is used as the indicator of stock status and is also used to determine catch limits in the MP.

4.2.2.3. LRP Role 3: Operational Control Points

When a reference point is used as an OCP in an MP (e.g., the provisional HCR in the PA Policy, DFO 2009), it is important to distinguish its role as a trigger for management action from that as the threshold in a measurable objective. To avoid confusion, alternative terms, e.g., “lower OCP” or simply “fishing cutoff” have been used to represent the role of the LRP as an OCP in a harvest control rule (Kronlund et al. 2018; Forrest et al. 2018).

4.3. CHANGE IN STOCK STATUS

Stock status estimates (biomass relative to the LRP: B/LRP) can change between assessments due to many factors, which include: acquisition of new or previously unavailable data; a change in analytical methods or model assumptions; or a change in the choice of reference points. Changes can be classified by source, as changes due to the monitored state (i.e., the numerator of B/LRP) or policy goal (i.e., the denominator of B/LRP) (Silvar-Viladomiu et al. 2021), or both the numerator and denominator through changes in analytical methods. Changes in stock status between assessments can be common. Silvar-Viladomiu et al. (2021) found reference points (based on F or biomass) changed between sequential assessments for 64% of 124 ICES stocks evaluated between 2011 to 2019. The two most important predictors for changes in status of biomass reference points were re-evaluation of the technical basis of the LRP and how assessment uncertainty was accounted for. There was no directional trend in the change in stock status that would represent a movement towards or away from sustainability (Silvar-Viladomiu et al. 2021).

Understanding and communicating the general causes for the change in status between assessments aligns with understanding and communicating the sources of uncertainty (Section 3.1) affecting the advice. Even slight changes in an estimate of a model parameter (e.g., h) could have significant implications on the estimated stock status. Inclusion of sources of error may result in more reliable estimates of stock status that may deviate from previous assessments.

Retrospective patterns are systematic changes in estimates of biomass or other model-estimated quantities that occur from the sequential addition or removal of a full year of data (Mohn 1999). Evaluation of potential retrospective patterns in stock assessment is a common way of diagnosing impacts of new data and possible structural error in the model. Temporal changes in model parameters (e.g., non-stationarity due to changes in productivity) is one cause of retrospective patterns that can result in model misspecification (e.g., Hurtado-Ferro et al. 2015; Szuwalski et al. 2018). The influence of retrospective on management advice can be explored using closed-loop simulation (Huynh et al. 2022).

The best available science and the guidance provided by DFO evolves over time. Changes in stock status due to changes in the LRP may reflect a change in methods (e.g., from a proportion of B_{MSY} to a proportion of B_0) or changes that reflect non-stationarity due to changes in productivity (e.g., regime shifts and long-term climate drift; ICES 2022). Changes in productivity can have a large influence on estimated reference points (e.g., B_{MSY}, B_0) due to change in parameters such as growth and recruitment (e.g., Barrett et al. 2022). The ICES approach for dealing with uncertainty in biological parameters (e.g., density dependence and changes in productivity) is to update reference points every benchmark (i.e., assessment framework, 3-5 years; ICES 2022); however, specific guidance from DFO varies from this. For example, the PA Policy (DFO 2009) suggests using the longest possible time series and to include high productivity periods in consideration of reference points. Conclusions from a DFO (2013) workshop on MSY reference points and the PA when productivity varies included “*changes in recruitment rates, natural mortality, fecundity, or growth rates are not considered to be* *appropriate reasons to change biomass limit reference points*". Guidance on defining reference points when there are changes in productivity is outside the scope of this paper (see Section 5.1), but changes in productivity may be an important explanatory factor to changes in stock status.

Marentette et al. (unpublished working paper) recommend that indicators and LRPs be re-evaluated when a new advisory framework (e.g., a new assessment model) or management paradigm (e.g., initiation of an MSE) is undertaken. The recommended approach to address changes in stock status is provided by Marentette et al. (unpublished working paper) and is based on clearly communicating the Science advice: (i) document the technical specifications and supporting rationales for the choice of indicator and LRP; (ii) document key uncertainties and assumptions; (iii) report stock status as the ratio of the indicator to LRP; and iv) evaluate the influence of various sources (e.g., additional data, model assumptions, technical basis for the LRP) on changes in stock status.

5. WHERE FURTHER WORK IS NEEDED

5.1. GUIDANCE ON TIME VARYING REFERENCE POINTS

Equilibrium reference points are influenced by non-stationary or regime shifts in key model productivity parameters such as M (Punt et al. 2021; O'Leary et al. 2020), growth (Barrett et al. 2022), and recruitment (A'mar 2009a,b; Maunder and Thorson 2019; Holt and Michielsens 2020). Equilibrium assumptions may also underlie many historical or empirical reference points. Management decisions based on static equilibrium reference points may not reflect stock dynamics in the future (Haltuch et al. 2009; O'Leary et al. 2020). Dynamic reference points have been proposed as one solution to address temporal changes in productivity parameters (Berger 2019). A dynamic B_0 can be defined as the biomass at any point in time that would have resulted if no fishing had occurred (MacCall et al. 1985). The dynamic B_0 time series can be estimated using the parameters of a stock assessment model and projecting the population forward over the same time period with $F = 0$ which makes an implicit assumption that temporal changes in biological parameters (e.g., M, weight-at-age, maturity-at-age, and recruitment) are independent of fishing and are not density-dependent (Berger 2019). These assumptions can be challenged and a continuous decrease in reference points over time due to a change in productivity seems in conflict with the objective of the definition of the LRP to avoid states of serious harm. Frequent changes to reference points may undermine trust in the management process (Walters 1998; Collie et al. 2012) and distract from more important components of the management process (Hilborn 2002). There is likely no one rule on when and how to use time-varying reference points, and their use will depend on the degree to which drivers of population dynamics are understood (Basson 1999; Haltuch and Punt 2011). Guidance on time-varying reference points is outside the scope of this paper but guidance is needed in the future.

5.2. POLICY GUIDANCE FOR LRPS BASED ON A PROPORTION OF B_0

The default policy guidance for LRPs in Canada is 0.4 B_{MSY}. In some situations, e.g., when h is high, 0.4 B_{MSY} can occur at a low level of depletion where other forms of serious harm may occur (Section 2). Other jurisdictions have set minimum LRPs in terms of B_0 (e.g., DAWR 2018: 0.2 B_0) or have made recommendations for considering setting a minimum LRP (ICES 2022, 0.1 B_0) for this reason. Given that reference points based on B_0 are commonly used in Canada (Marentette et al. unpublished working paper), default policy guidance for LRPs based on B_0 is recommended for future consideration in Canadian harvest strategy policies (in addition to the existing guidance of 0.4 B_{MSY}).

6. CONCLUSIONS

To support the process of choosing LRPs and estimating stock status, Marentette et al. (unpublished working paper) identified the need for technical guidance on various aspects of estimating reference points, considerations for selecting candidate indicators of stock status and LRPs, and options for reporting a single stock status (Section 1.2). These were addressed herein and the following recommendations are provided:

- Evaluate several candidate LRPs and look for agreements/disagreements among methods (van Deurs et al. 2021; ICES 2022).
- When the functional form of the SRR and/or estimates of *h* are highly uncertain, consider defining an LRP:
  - using a proxy for B_{MSY} (ICES 2022);
  - using an ensemble from multiple models that capture model (structural) uncertainty.
- When *h* is assumed to be high, LRPs based on a proportion of B_{MSY} may be below the minimum observed historical biomass or below a threshold for other sources of serious harm so consider:
  - setting a minimum LRP in terms of B_0 (e.g., 0.2 B_0, DAWR 2018; 0.1 B_0, MF 2008; ICES 2022) to avoid low depletion or the potential for Allee effects (e.g., at a threshold of 0.15-0.25 B_0, Perälä and Kuparinen 2017; Perälä et al. 2022);
  - using a proxy for B_{MSY} (e.g., B_0 or equilibrium biomass at $F_{X\%SPR}$); or
  - using a historical biomass (e.g., $B_{recover}$, B_{loss}) or another threshold that represents an undesirable state.
- Consider using closed-loop simulation with multiple OMs to capture a range of uncertainty in stock and fishery dynamics, incorporate feedback between MPs and stock outcomes, and provide harvest advice in compliance with the PA Policy and the FSP.
- There are many data-limited approaches to estimating LRPs and stock status, all of which require assumptions that should be given careful consideration depending on the stock life-history, fishing history, and data-availability.
- Catch-only methods are generally not recommended for assessing individual stock status, but can provide a preliminary source of information in data-limited situations.
- For data-limited stocks, consider using closed-loop simulation methods to evaluate performance of alternative data-limited MPs over a range of uncertainties in stock and fishery dynamics.
- When multiple models (assessment models or OMs) can be equally weighted or plausibility weights have been determined, these weights can be used in an ensemble approach to estimate stock status while capturing model uncertainty.
- If plausibility weights are not available for multiple models (e.g., IWC approach of estimating performance using closed-loop simulation), an indicator of stock status can be selected from a single base-case or most pessimistic model (understanding that model uncertainty is not captured) or using an empirical indicator.
- LRPs should be re-evaluated during each advisory framework review.
- Clearly communicate the advice on stock status and the LRP so sources of any changes in stock status between assessments can be understood.

7. REFERENCES CITED

A’mar, Z.T., Punt, A.E., and Dorn, M.W. 2009a. The impact of regime shifts on the performance of management strategies for the Gulf of Alaska walleye pollock (*Theragra chalcogramma*) fishery. Canadian Journal of Fisheries and Aquatic Sciences, 66: 2222-2242.

A’mar, Z.T., Punt, A.E., and Dorn, M.W. 2009b. The evaluation of two management strategies for the Gulf of Alaska walleye pollock fishery under climate change. ICES Journal of Marine Science, 66: 1614-1632.

Agnew, D.J., Gutiérrez, N.L., Stern-Pirlot, A., and Hoggarth, D.D. 2014. The MSC experience: developing an operational certification standard and a market incentive to improve fishery sustainability. ICES Journal of Marine Science, 71: 216-225.

Albertsen, C.M., and Trijoulet, V. 2020. Model-based estimates of reference points in an age-based state-space stock assessment model. Fisheries Research, 230: 105618.

Anderson, S.C., Branch, T.A., Ricard, D., and Lotze, H.K., 2012. Assessing global marine fishery status with a revised dynamic catch-based method and stock-assessment reference points. ICES Journal of Marine Science, 69: 1491-1500.

Anderson, S.C., Cooper, A.B., Jensen, O.P., Minto, C., Thorson, J.T., Walsh, J.C., Afflerbach, J., Dickey-Collas, M., Kleisner, K.M., Longo, C., Osio, G.C., Ovando, D., Mosqueira, I., Rosenberg, A.A., and Selig, E.R. 2017. Improving estimates of population status and trend with superensemble models. Fish and Fisheries, 18: 732-741.

Anderson, S.C., Forrest, R.E., Huynh, Q.C., and Keppel, E.A. 2021. A management procedure framework for groundfish in British Columbia. DFO Can. Sci. Advis. Sec. Res. Doc. 2021/007.

Barrett, T.J., Hordyk, A.R., Barrett, M.A., and van den Heuvel, M.R. 2022. Spatial and temporal differences in fecundity of Atlantic herring (*Clupea harengus*) off Nova Scotia and consequences for biological reference points. Canadian Journal of Fisheries and Aquatic Sciences, 79: 1086-1096.

Barrowman, N.J., and Myers, R.A. 2000. Still more spawner-recruitment curves: the hockey stick and its generalizations. Canadian Journal of Fisheries and Aquatic Sciences, 57: 665-676.

Basson, M. 1999. The importance of environmental factors in the design of management procedures. ICES Journal of Marine Science, 56: 933-942.

Berger, A.M. 2019. Character of temporal variability in stock productivity influences the utility of dynamic reference points. Fisheries Research, 217: 185-197.

Berkson, J., and Thorson, J.T. 2015. The determination of data-poor catch limits in the United States: is there a better way? ICES Journal of Marine Science, 72: 237-242.

Berkson, J., Barbieri, L., Cadrin, S., Cass-Calay, S.L., Crone, P., Dorn, M., Friess, C., Kobayashi, D., Miller, T.J., Patrick, W.S., Pautzke, S., Ralston, S., and Trianni, M. 2011. Calculating acceptable biological catch for stocks that have reliable catch data only (Only reliable catch stocks – ORCS). NOAA Tech. Memo. 1–56 NMFS-SEFSC-616.

Beverton, R.J.H., and Holt, S.J. 1957. On the dynamics of exploited fish populations. Chapman and Hall, London, Fish and Fisheries Series No. 11, Reprint 1993.

Bouch, P., Minto, C., and Reid, D.G. 2021. Comparative performance of data-poor CMSY and data-moderate SPiCT stock assessment methods when applied to data-rich, real-world stocks. ICES Journal of Marine Science, 78(1): 264-276.

Braccini, M., Brooks, E.N., Wise, B., and McAuley, R. 2015. Displaying uncertainty in the biological reference points of sharks. Ocean and Coastal Management, 116: 143-149.

Branch, T.A., Jensen, O.P., Ricard, D., Ye, Y., and Hilborn, R. 2011. Contrasting global trends in marine fishery status obtained from catches and from stock assessments. Conservation Biology, 25: 777-786.

Brooks, E.N. 2013. Effects of variable reproductive potential on reference points for fisheries management. Fisheries Research, 138: 152-158.

Brooks, E.N., and Deroba, J.J. 2015. When “data” are not data: the pitfalls of post hoc analyses that use stock assessment model output. Canadian Journal of Fisheries and Aquatic Sciences, 72: 634-641.

Brosset, P., Doniol-Valcroze, T., Swain, D.P., Lehoux, C., Van Beveren, E., Mbaye, B.C., Emond, K., and Plourde, S. 2019. Environmental variability controls recruitment but with different drivers among spawning components in Gulf of St. Lawrence herring stocks. Fisheries Oceanography, 28: 1-17.

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

Butterworth, D.S., Punt, A.E., and Smith, A.D.M. 1996. On plausible hypotheses and their weighting, with implications for selection between variants of the Revised Management Procedure. Reports of the International Whaling Commission, 46: 637-640.

Caddy, J.F. 2002. Limit reference points, traffic lights, and holistic approaches to fisheries management with minimal stock assessment input. Fisheries Research, 56: 133-137.

Cadigan, N.G. 2013. Fitting a non-parametric stock-recruitment model in R that is useful for deriving MSY reference points and accounting for model uncertainty. ICES Journal of Marine Science, 70: 56-67.

Cadrin, S.X. 2012. Unintended consequences of MSY proxies for defining overfishing. ICES Evolution of management frameworks to prevent overfishing. ICES CM 2012/L:23.

Carruthers, T.R., and Hordyk, A.R. 2018. The Data-Limited Methods Toolkit (DLMtool): An R package for informing management of data-limited populations. Methods in Ecology and Evolution, 9(12): 2388-2395.

Carruthers, T.R., Walters, C.J., and McAllister, M.K. 2012. Evaluating methods that classify fisheries stock status using only fisheries catch data. Fisheries Research, 119-120: 66-79.

Carruthers, T.R., Punt, A.E., Walters, C.J., MacCall, A., McAllister, M.K., Dick, E.J., and Cope, J. 2014. Evaluating methods for setting catch limits in data-limited fisheries. Fisheries Research, 153: 48-68.

Carruthers, T.R., and Hordyk, A.R. 2019. Using management strategy evaluation to establish indicators of changing fisheries. Canadian Journal of Fisheries and Aquatic Sciences, 76: 1653-1668.

Carruthers, T.R., Kell, L.T., Butterworth, D.D., Maunder, M.N., Geromont, H.F., Walters, C., McAllister M.K. et al. 2016. Performance review of simple management procedures. ICES Journal of Marine Science, 73: 464-482.

Chaput, G. 2015. Considerations for defining reference points for Atlantic salmon that conform to the precautionary approach. DFO Can. Sci. Advis. Sec. Res. Doc. 2015/074.

Chrysafi, A., and Cope, J.M. 2019. Testing approaches to determine relative stock abundance priors when setting catch recommendations using data-limited methods. Fisheries Research, 219: 105343.

Chrysafi, A., Cope, J.M., and Kuparinen, A. 2019. Eliciting expert knowledge to inform stock status for data-limited stock assessments. Marine Policy, 101: 167-176.

Clark, W.G. 1991. Groundfish exploitation rates based on life history parameters. Canadian Journal of Fisheries and Aquatic Sciences, 48, 734–750.

Clark, W.G. 2002. F35% revisited ten years later. North American Journal of Fisheries Management, 22: 251-257.

Collie, J.S., Peterman, R.M., and Zuehlke, B.M. 2012. A fisheries risk-assessment framework to evaluate trade-offs among management options in the presence of time-varying productivity. Canadian Journal of Fisheries and Aquatic Sciences, 69, 209-223.

Cook, A.M., Hubley, P.B., Denton, C., and Howse, V. 2020. 2018 Framework Assessment of the American Lobster (*Homarus americanus*) in LFA 27-33. DFO Can. Sci. Advis. Sec. Res. Doc. 2020/017.

Cope, J.M. 2013. Implementing a statistical catch-at-age model (Stock Synthesis) as a tool for deriving overfishing limits in data-limited situations. Fisheries Research, 142: 3-14.

Cope, J.M., and Punt, A.E. 2009. Length-based reference points for data-limited situations: Applications and restrictions. Marine and Coastal Fisheries: Dynamics, Management and Ecosystem Science, 1(1): 169-186.

Costello, C., Ovando, D., Hilborn, R., Gaines, S.D., Deschenes, O., and Lester, S.E. 2012. Status and solutions for the world's unassessed fisheries. Science, 338: 517-520.

DAWR (Department of Agriculture and Water Resources, Australian Government). 2018. Guidelines for the Implementation of the Commonwealth Fisheries Harvest Strategy Policy. 2nd edition. Canberra, June. CC BY 4.0. 42 p.

de la Mare, W.K. 1998. Tidier fisheries management requires a new MOP (management oriented paradigm). Reviews in Fish Biology and Fisheries, 8: 349-356.

Deriso, R.B. 1980. Harvesting strategies and parameter estimation for an age-structured model. Canadian Journal of Fisheries and Aquatic Sciences, 37: 268-282

DFO (Fisheries and Oceans Canada). 2002. National workshop on reference points for gadoids. DFO Can. Sci. Advis. Sec. Proc. Ser. 2002/033.

DFO. 2004. Proceedings of the national meeting on applying the precautionary approach in fisheries management; February 10-12, 2004. Can. Sci. Advis. Sec. Proc. Ser. 2004/003.

DFO. 2006. A harvest strategy compliant with the precautionary approach. DFO Can. Sci. Advis. Sec. Sci. Advis. Rep. 2006/023.

DFO. 2008. Proceedings of the precautionary approach workshop on shrimp and prawn stocks and fisheries. DFO Can. Sci. Advis. Sec. Proceed. Ser. 2008/031.

DFO. 2009a. A fishery decision-making framework incorporating the precautionary approach. Last updated 2009-03-23.

DFO. 2011. Western component (4Xopqrs5) pollock management strategy evaluation. DFO Can. Sci. Advis. Sec. Sci. Advis. Rep. 2011/054.

DFO. 2014. Stock assessment and harvest advice for rock sole (*Lepidopsetta* spp.) in British Columbia. DFO Can. Sci. Advis. Sec. Sci. Advis. Rep. 2014/039.

DFO. 2015. Development of reference points for Atlantic salmon (*Salmo salar*) that conform to the Precautionary Approach. DFO Can. Sci. Advis. Sec. Sci. Advis. Rep. 2015/058.

DFO. 2016. Proceedings of the national peer review on the development of technical guidelines for the provision of scientific advice on the various elements of Fisheries and Oceans Canada precautionary approach framework; February 28-March 1, 2012. DFO Can. Sci. Advis. Sec. Proceed. Ser. 2015/005.

DFO. 2018a. Units 1+2 redfish management strategy evaluation. DFO Can. Sci. Advis. Sec. Sci. Advis. Rep. 2018/033.

DFO. 2018b. Limit reference points for Atlantic salmon rivers in DFO Gulf region. DFO Can. Sci. Advis. Sec. Sci. Advis. Rep. 2018/015.

DFO. 2019a. Guidance for the development of rebuilding plans under the precautionary approach framework: growing stocks out of the critical zone.

DFO. 2019b. Assessment of British Columbia Pacific cod for Areas 3CD, and 5ABCD in 2018. DFO Can. Sci. Advis. Sec. Sci. Advis. Rep. 2019/008.

DFO. 2020a. Evaluation of management procedures for the inside population of yelloweye rockfish rebuilding plan in British Columbia. DFO Can. Sci. Advis. Sec. Sci. Advis. Rep. 2020/056.

DFO. 2020b. Evaluating the robustness of candidate management procedures in the BC sablefish (*Anoplopoma fibria*) fishery for 2019-2020. DFO Can. Sci. Advis. Sec. Sci. Advis. Rep. 2020/025.

DFO. 2020c. Evaluation of potential rebuilding strategies for the robustness of outside yelloweye rockfish in British Columbia. DFO Can. Sci. Advis. Sec. Sci. Advis. Rep. 2020/024.

DFO. 2021. Stock status update with application of management procedures for Pacific herring (*Clupea pallasii*) in British Columbia. DFO Can. Sci. Advis. Sec. Sci. Advis. Rep. 2021/039.

DFO. 2022a. Methodologies and guidelines for defining limit reference points for Pacific salmon. DFO Can. Sci. Advis. Sec. Sci. Advis. Rep. 2022/030.

DFO. 2022b. Stock assessment of herring in Northwest Atlantic Fisheries Organization (NAFO) fishing areas 4VWX. DFO Can. Sci. Advis. Sec. Sci. Advis. Rep. 2022/050.

DFO. 2023a. Science advice on guidance for limit reference points under the fish stocks provisions. DFO Can. Sci. Advis. Sec. Sci. Advis. Rep. 2023/009.

DFO. 2023b. Stock status update of Atlantic halibut (*Hippoglossus hippoglossus*) on the Scotian Shelf and Southern Grand Banks in NAFO Divisions 3NOPS4VWX5ZC for 2022. DFO Can. Sci. Advis. Sec. Sci. Advis. Rep. 2023/020.

DFO. 2024. Rapid status approximations for Pacific Salmon derived from integrated expert assessments under DFO's Wild Salmon Policy. DFO Can. Sci. Advis. Sec. Sci. Resp. In press.

Dietterich, T.G. 2000. Ensemble methods in machine learning. In Multiple classifier systems. Springer, Berlin Heidelberg. pp. 1-15.

Dorn, M.W. 2002. Advice on west coast rockfish harvest rates from Bayesian meta-analysis of stock-recruit relationships. North American Journal of Fisheries Management, 22: 280-300.

Dowling, N.A., Dichmont, C.M., Haddon, M., Smith, D.C., Smith, A.D.M., and Sainsbury, K. 2015. Empirical harvest strategies for data-poor fisheries: a review of the literature. Fisheries Research, 171: 141-153.

Drescher, M., Perera, A.H., Johnson, C.J., Buse, L.J., Drew, C.A., and Burgman, M.A. 2013. Toward rigorous use of expert knowledge in ecological research. Ecosphere, 4: 1-26.

eCFR (Electronic Code of Federal Regulations) 2021. Title 50: Wildlife and Fisheries, Part 600: Magnuson-Stevens Act Provisions, Subpart D: National Standards.

EU (European Union). 2013. European Union: Common fisheries policy.

FAO (Food and Agriculture Organization of the United Nations). 1995. Code of Conduct for responsible Fisheries. Rome, FAO.

FAO. 2001. BAYES-SA, Bayesian stock assessment methods in fisheries management – User manual. FAO Corporate Document Repository Y1958E.

Fletcher, R.I. 1978. On the restructuring of the Pella-Tomlinson system. Fishery Bulletin, 76: 515-521.

Flood, M.J., Stobutzki, I., Andrews, J., Ashby, C., Begg, G.A., Fletcher, R., Gardner, C., Georgeson, L., Hansen, S., Hartmann, K. and Hone, P. 2016. Multijurisdictional fisheries performance reporting: how Australia's nationally standardised approach to assessing stock status compares. Fisheries Research, 183: 559-573.

Forrest, R.E., McAllister, M.K., Dorn, M., Martell, S.J.D, Stanley, R. 2010. Hierarchical Bayesian estimation of productivity and reference points for Pacific rockfishes (*Sebastes* spp.) under alternative assumptions about the stock-recruit function. Canadian Journal of Fisheries and Aquatic Sciences, 67: 1611-1634.

Forrest, R.E., Holt, K.E, and Kronlund, A.R. 2018. Performance of alternative harvest control rules for two Pacific groundfish stocks with uncertain natural mortality: Bias, robustness and trade-offs. Fisheries Research, 206: 259-286.

Forrest, R.E., Anderson, S.C., Grandin, C.J., and Starr, P.J. 2020. Assessment of Pacific cod (*Gadus macrocephalus*) for Hecate Strait and Queen Charlotte Sound (Area 5ABCD), and West Coast Vancouver Island (Area 3CD) in 2018. DFO Can. Sci. Advis. Sec. Res. Doc. 2020/070.

Fox, W.W., Jr. 1970. An exponential surplus-yield model for optimizing exploited fish populations. Transactions of the American Fisheries Society, 99: 80-88.

Francis, R.C. 1974. Relationship of fishing mortality to natural mortality at the level of maximum sustainable yield under the logistic stock production model. Journal of the Fisheries Research Board of Canada, 31(9):1539-1542.

Free, C.M., Jensen, O.P., Wiedenmann, J., and Deroba, J.J. 2017. The refined ORCS approach: A catch-based method for estimating stock status and catch limits for data-poor fish stocks. Fisheries Research, 193: 60-70.

Free, C.M., Jensen, O.P., Anderson, S.C., Gutierrez, N.L., Kleisner, K.M., Longo, C., Minto, C., Osio, G.C., and Walsh, J.C. 2020. Blood from a stone: Performance of catch-only methods in estimating stock biomass status. Fisheries Research, 223: 105452.

Froese, R. 2004. Keep it simple: three indicators to deal with overfishing. Fish and Fisheries, 5: 86-91.

Froese, R., and Kesner-Reyes, K. 2002. Impact of fishing on the abundance of marine species. ICES CM 12, 1–12 2002/L.

Froese, R. and Pauly, D. 2022. FishBase. Version (02/2022).

Froese, R., Coro, G., Kleisner, K., and Demirel, N. 2016. Revisiting safe biological limits in fisheries. Fish and Fisheries, 17: 193-209.

Froese, R., Demirel, N., Coro, G., Kleisner, K. M., and Winker, H. 2017. Estimating fisheries reference points from catch and resilience. Fish and Fisheries, 18, 506-526.

Froese, R., Winker, H., Coro, G., Demirel, N., Tsikliras, A.C., Dimarchopoulou, D., Scarcella, G., Probst, W.N., Dureuil, M., and Pauly, D. 2018. A new approach for estimating stock status from length frequency data. ICES Journal of Marine Science, 75: 2004-2015.

Froese, R., Winker, H., Coro, G., Palomares, M.L., Tsikliras, A.C., Dimarchopoulou, D., Touloumis, K., Demirel, N., Vianna, G.M.S., Scarcella, G., Schijns, R., Liang, C., and Pauly, D. 2022. Catch time series as the basis for fish stock assessments: the CMSY++ method. Added to the GEOMAR repository January 12, 2022. [Accessed December 13, 2022].

Gabriel, W.L., and Mace, P.M. 1999. A review of biological reference points in the context of the precautionary approach. In: Proceedings of the fifth national NMFS stock assessment workshop: providing scientific advice to implement the precautionary approach under the Magnuson-Stevens fishery conservation and management act. NOAA Tech Memo NMFSF/SPO-40. pp. 34-45.

Gabriel, W.L., Sissenwine, M.P., and Overholtz, W.J. 1989. Analysis of spawning stock biomass per recruit: an example for Georges Bank haddock. North American Journal of Fisheries Management, 9: 383-391.

Geromont, H.F., and Butterworth, D.D. 2015. Generic management procedures for data-poor fisheries: forecasting with few data. ICES Journal of Marine Science, 72: 251-261.

Goodyear, C.P. 1993. Spawning stock biomass per recruit in fisheries management: foundation and current use. In: Risk evaluation and biological reference points for fisheries management. Eds: Smith, S.J., Hunt, J.J., and Rivard, D. Canadian Special Publication of Fisheries and Aquatic Sciences, 120: 67-82.

Grant, S.C.H., and Pestal, G. 2012. Integrated biological status assessments under the Wild Salmon Policy using standardized metrics and expert judgement: Fraser River sockeye salmon (*Oncorhynchus nerka*) case studies. Can. Sci. Advis. Sec. Res. Doc. 2012/106: v + 132 pp.

Gulland, J.A., and Boerema, L.K. 1973. Scientific advice on catch levels. Fish Bulletin, 71: 325-335.

Haddon, M. 2012. Reducing Uncertainty in Stock Status: Harvest Strategy Testing, Evaluation, and Development. General Discussion and Summary. CSIRO Marine and Atmospheric Research. 42 pp.

Haggarty D.R., Huynh, Q.C., Forrest, R.E., Anderson, S.C., Bresch, M.J., and Keppel, E.A. 2022. Evaluation of potential rebuilding strategies for inside yelloweye rockfish (*Sebastes ruberrimus*) in British Columbia. DFO Can. Sci. Advis. Sec. Res. Doc. 2021/008. vi + 139p.

Halliday, R.G., Fanning, L.P., and Mohn, R.K. 2001. Use of the traffic light method in fishery management planning. DFO Can. Sci. Advis. Sec. Res. Doc. 2001/108. vii + 42 p.

Haltuch, M.A, and Punt, A.E. 2011. The promises and pitfalls of including decadal-scale climate forcing of recruitment in groundfish stock assessment. Canadian Journal of Fisheries and Aquatic Sciences, 68: 912-926.

Haltuch, M.A., Punt, A.E., and Dorn, M.W. 2008. Evaluating alternative estimators of fishery management reference points. Fisheries Research, 94: 290-303.

Haltuch, M.A., Punt, A.E., and Dorn, M.W. 2009. Evaluating the estimation of fishery management reference points in a variable environment. Fisheries Research, 100: 42-56.

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: 62-69.

Harford, W.J., Amoroso, R., Bell, R.J., Caillaux, M., Cope, J.M., Dougherty, D., et al. 2021. Multi-indicator harvest strategies for data-limited fisheries: a practitioner guide to learning and design. Frontiers in Marine Science, 8: 757877.

He, X., Ralston, S., and MacCall, A.D. 2011. Interactions of age-dependent mortality and selectivity functions in age-based stock assessment models. Fishery Bulletin, 109: 198-216.

Health Canada. 2018. Weight of evidence: general principles and current applications at Health Canada. Prepared for the Task Force on Scientific Risk Assessment by the Weight of Evidence Working Group. 18 pp.

Hilborn, R. 1985. Simplified calculation of optimum spawning stock size from Ricker's stock recruitment curve. Canadian Journal of Fisheries and Aquatic Sciences, 42: 1833-1834.

Hilborn, R. 2002. The dark side of reference points. Bulletin of Marine Science, 70(2): 403-408.

Hilborn, R., and Walters, C.J. 1992. Quantitative fisheries stock assessment: choice, dynamics and uncertainty. Chapman and Hall, New York. 570 pp.

Holt, C.A., and Michielsens, C.G.J. 2020. Impact of time-varying productivity on estimated stock–recruitment parameters and biological reference points. Canadian Journal of Fisheries and Aquatic Sciences, 77: 836-847.

Holt, C.A., Cass, A., Holtby, B., and Riddell, B. 2009. Indicators of status and benchmarks for conservation units in Canada's wild salmon policy. DFO Can. Sci. Advis. Res. Doc. 2009/058.

Holt, K.R., Starr, P.J., Haigh, R., and Krishka, B. 2016. Stock assessment and harvest advice for Rock Sole (*Lepidopsetta* spp.) in British Columbia. DFO Can. Sci. Advis. Sec. Res. Doc. 2016/009. ix+256p.

Holt, C.A., Holt, K., Warkentin, L., Wor, C., Connors, B., Grant, S., Huang, A.-M., and Marentette, J. 2023a. Guidelines for defining limit reference points for Pacific salmon stock management units. DFO Can. Sci. Advis. Sec. Res. Doc. 2023/009.

Holt, K.R., Holt, C.A., Warkentin, L., Wor, C., Davis, B., Arbeider, M., Bokvist, J., Crowley, S., Grant, S., Luedke, W., McHugh, D., Picco, C., and Van Will, P. 2023b. Case study applications of LRP estimation methods to Pacific salmon stock management units. DFO Can. Sci. Advis. Sec. Res. Doc. 2023/010.

Hordyk, A.R. 2021. LBSPR: An R package for simulation and estimation using life-history ratios and length composition data. [Accessed May 22, 2022].

Hordyk, A.R., and Carruthers, T.R. 2018. A quantitative evaluation of a qualitative risk assessment framework: Examining the assumptions and predictions of the Productivity Susceptibility Analysis (PSA). PLOS ONE 13: e0198298.

Hordyk, A., Ono, K., Valencia, S., Loneragan, N., and Prince, J. 2015a. A novel length-based empirical estimation method of spawning potential ratio (SPR), and tests of its performance, for small-scale, data-poor fisheries. ICES Journal of Marine Science, 72: 217-231.

Hordyk, A.R., Loneragan, N.R., and Prince, J.D. 2015b. An evaluation of an iterative harvest strategy for data-poor fisheries using the length-based spawning potential ratio assessment methodology. Fisheries Research, 171: 20-32.

Hordyk, A.R., Prince, J.D., Carruthers, T.R., and Walters, C.J. 2019. Comment on "A new approach for estimating stock status from length frequency data" by Froese et al. (2018). ICES Journal of Marine Science, 76: 457-460.

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

Huynh, Q.C., Legault, C.M., Hordyk, A.R., and Carruthers, T.R. 2022. A closed-loop simulation framework and indicator approach for evaluating impacts of retrospective patterns in stock assessments. ICES Journal of Marine Science, 79: 2003-2016.

ICES (International Council for the Exploration of the Sea). 1988. Reports of the ICES Advisory Committee on Fishery Management, 1987. ICES Co-operative Research Report, 153. 415 pp.

ICES. 2013. Report of the Workshop on guidelines for management strategy evaluations (WKGMSE). 21-23 January 2013, Copenhagen, Denmark. ICES CM 2013/ACOM:39. 122 pp.

ICES. 2015. Report of the benchmark workshop on North Sea stocks (WKNSEA). 2-6 February 2015, Copenhagen, Denmark. ICES CM 2015/ACOM:32. 253 pp.

ICES. 2018. 16.4.3.2. ICES reference points for stocks in categories 3 and 4. ICES Technical Guidelines, published 13 February 2018. 50 pp.

ICES. 2020. The third workshop on guidelines for management strategy evaluations (WKGMSE3). ICES Scientific Reports. 2:116. 112 pp.

ICES. 2021a. 16.4.3.1. ICES fisheries management reference points for category 1 and 2 stocks. ICES Technical Guidelines, published 1 March 2021. 19 pp.

ICES. 2021b. Advice on fishing opportunities. In: Report of the ICES Advisory Committee, 2021. ICES Advice 2021, section 1.1.1, published 21 July 2021. 9 pp.

ICES. 2022. Workshop on ICES reference points (WKREF1). ICES Scientific Reports. 4:2. 70 pp.

Jardim, E., Azevedo, M., Brodziak, J., Brooks, E.N., Johnson, K.F., Klibansky, N., Millar, C.P., Minto, C., Mosqueira, I., Nash, R.D.M., Vasilakopoulos, P., and Wells, B.K. 2021. Operationalizing ensemble models for scientific advice to fisheries management. ICES Journal of Marine Science, 78: 1209-1216.

Johnson, K.F., Monnahan, C.C., McGilliard, C.R., Vert-Pre, K.A., Anderson, S.C., Cunningham, C.F., et al. 2015. Time-varying natural mortality in fisheries stock assessment models: identifying a default approach. ICES Journal of Marine Science, 72: 137-150.

Kell, L.T., De Oliveira, J.A.A., Punt, A.E., McAllister, M.K., and Kuikka, S. 2006. Operational management procedures: an introduction to the use of evaluation frameworks. Developments in Aquaculture and Fisheries Science, 36: 379-407.

Kleisner, K., and Pauly, D. 2011. Stock catch status plots of fisheries for regional seas. In: The state of biodiversity and fisheries in regional seas. Fisheries Centre Research Reports. Eds: Christensen, V., Lai, S., and Palomares, M. University of British Columbia, Vancouver, Canada, pp. 37-40.

Kleisner, K., Zeller, D., Froese, R., and Pauly, D. 2013. Using global catch data for inference on the world's marine fisheries. Fish and Fisheries, 14: 293-311.

Kronlund, A.R., Forrest, R.E., Cleary, J.S., Grinnell, M.H. 2018. The selection and role of limit reference points for Pacific herring (*Clupea pallasii*) in British Columbia, Canada. DFO Can. Sci. Advis. Sec. Res. Doc. 2018/009. ix + 125 p.

Kronlund, A.R., Marentette, J.R., Olmstead, M., Shaw, J., and Beauchamp, B. 2021. Considerations for the design of rebuilding strategies for Canadian fish stocks. DFO Can. Sci. Advis. Sec. Res. Doc. 2021/051. vii + 134 p.

Larcombe, J., Noriega, R., and Stobutzi, I. 2015. Reducing uncertainty in fisheries stock status. Australian Bureau of Agricultural and Resource Economics and Sciences, Canberra, xi+220 pp

Link, W.A. 1999. Modeling pattern in collections of parameters. Journal of Wildlife Management, 63: 1017-1027.

MacCall, A.D. 1990. Dynamic Geography of Marine Fish Populations. University of Washington Sea Grant Program, Seattle. 158 pp.

MacCall, A.D., Klingbeil, R.A., and Methot, R.D. 1985. Recent increased abundance and potential productivity of Pacific mackerel (*Scomber japonicus*). California Cooperative Oceanic Fisheries Investigations Report, 26: 119-129.

Mace, P.M. 1994. Relationships between common biological reference points used as thresholds and targets of fisheries management strategies. Canadian Journal of Fisheries and Aquatic Sciences, 51: 110-122.

Mace, P.M. 2001. A new role for MSY in single-species and ecosystem approaches to fisheries stock assessment and management. Fish and Fisheries, 2: 2-32.

Mace, P.M., and Doonan, I.J. 1988. A generalised bioeconomic simulation model for fish population dynamics. New Zealand Fishery Assessment Research Document 88/4. Fisheries Research Centre, MAFFish, POB 297, Wellington, NZ.

Mace, P.M., and Sissenwine, M.P. 1993. How much spawning per recruit is enough? *In* S.J. Smith, J.J. Hunt and D. Rivard (ed.). Risk evaluation and biological reference points for fisheries management. Can. Spec. Publ. Fish. Aquat. Sci. 120: 101-118.

Magnusson, A., and Hilborn, R. 2007. What makes fisheries data informative? Fish and Fisheries, 8: 337-358.

Magnusson, A., Punt, A.E., and Hilborn, R. 2013. Measuring uncertainty in fisheries stock assessment: the delta method, bootstrap, and MCMC. Fish and Fisheries, 14: 325-342.

Mangel, M., Brodziak, J., and DiNard, G. 2010. Reproductive ecology and scientific inference of steepness: a fundamental metric of population dynamics and strategic fisheries management. Fish and Fisheries, 11: 89-104.

Mangel, M., MacCall, A.D., Brodziak, J., Dick, E.J., Forrest, R.E., Pourzand, R., and Ralston, S. 2013. A perspective on steepness, reference points, and stock assessment. Canadian Journal of Fisheries and Aquatic Sciences, 70: 930-940.

Martell, S., and Froese, R. 2013. A simple method for estimating MSY from catch and resilience. Fish and Fisheries, 14: 504-515.

Maunder, M.N., and Punt, A.E. 2013. A review of integrated analysis in fisheries stock assessment. Fisheries Research, 142: 61-74.

Maunder, M.N., and Thorson, J.T. 2019. Modeling temporal variation in recruitment in fisheries stock assessment: A review of theory and practice. Fisheries Research, 217: 71-86.

Maunder, M.N., Xu, H., Lennert-Cody, C.E., Valero, J.L., Aires-da-Silva, A., and Minte-Vera, C. 2020. Implementing reference point-based fishery harvest control rules within a probabilistic framework that considers multiple hypotheses. Inter-American Tropical Tuna Commission. 11th Meeting. 54pp.

McAllister, M.K., Pikitch, E.K., and Babcock, E.A. 2001. Using demographic methods to construct Bayesian priors for the intrinsic rate of increase in the Schaefer model and implications for stock rebuilding. Canadian Journal of Fisheries and Aquatic Sciences, 58: 1871-1890.

Mesnil, B., and Rochet, M.J. 2010. A continuous hockey stick stock–recruit model for estimating MSY reference points. ICES Journal of Marine Science, 67: 1780-1784.

MF (New Zealand Government, Ministry of Fisheries). 2008. Harvest Strategy Standard for New Zealand Fisheries. October 2008. 30p.

MF. 2011. Operational Guidelines for New Zealand's Harvest Strategy Standard, Revision 1, June 2011. 78 p.

Michielsens, C.G.J., and McAllister, M.K. 2004. A Bayesian hierarchical analysis of stock–recruit data: quantifying structural and parameter uncertainties. Canadian Journal of Fisheries and Aquatic Sciences, 61: 1032-1047.

Miller, D.C.M., and Shelton, P.A. 2010. "Satisficing" and trade-offs: Evaluating rebuilding strategies for Greenland halibut off the east coast of Canada. ICES Journal of Marine Science, 67: 1896-1902.

Miller, T.J., and Brooks, E.N. 2020. Steepness is a slippery slope. Fish and Fisheries, 22: 634-645.

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.

MSC (Marine Stewardship Council) 2018. MSC Fisheries Standard. Version 2.01, 31 August 2018.

Murray, L.G., and Seed, R. 2010. Determining whether catch per unit effort is a suitable proxy for relative crab abundance. Marine Ecology Progress Series, 401: 173-182.

Musick, J.A. 1999. Criteria to define extinction risk in marine fishes. Fisheries, 24: 6-14.

Myers, R.A., and Mertz, G. 1998. The limits of exploitation: a precautionary approach. Ecological Applications, 8 Supplement, s165-s169.

Myers, R.A., Bowen, K.G., and Barrowman, N.J. 1999. Maximum reproductive rate of fish at low population sizes. Canadian Journal of Fisheries and Aquatic Sciences, 56: 2404-2419.

Myers, R.A., Rosenberg, A.A., Mace, P.M., Barrowman, N., and Restrepo, V.R. 1994. In search of thresholds for recruitment overfishing. ICES Journal of Marine Science, 51: 191-205

Myers, R.A., Barrowman, N.J., Hutchings, J.A., and Rosenberg, A.A. 1995a. Population dynamics of exploited fish stocks at low population levels. Science, 269: 1106-1108.

Myers, R.A., Bridson J., and Barrowman, N.J. 1995b. Summary of world-wide stock and recruitment data. Canadian Technical Reports of Fisheries and Aquatic Sciences 2024. iv+ 327p.

Myers, R.A., Barrowman, B., Hilborn, R., and Kehler, D. 2002. Inferring Bayesian priors with limited direct data: applications to risk analysis. North American Journal of Fisheries Management, 22: 351-364.

NAFO (Northwest Atlantic Fisheries Organization). 2004. Report of the NAFO Study Group on Limit Reference Points. NAFO SCS Doc. 04/12, Serial No. N4980.

O'Connell, M.F., Dempson, J.B., and Chaput, G. 2006. Aspects of the life history, biology, and population dynamics of Atlantic salmon (*Salmo salar* L.) in Eastern Canada. DFO Can. Sci. Advis. Sec. Res. Doc. 2006/014.

O'Leary, C.A., Thorson, J.T., Miller, T.J., and Nye, J.A. 2020. Comparison of multiple approaches to calculate time-varying biological reference points in climate-linked population-dynamics models. ICES Journal of Marine Science, 77: 930-941.

Ovando, D., Hilborn, R., Monnahan, C., Rudd, M., Sharma, R., Thorson, J.T., Rousseau, Y., and Ye, Y. 2021. Improving estimates of the state of global fisheries depends on better data. Fish and Fisheries, 22: 1377-1391.

Ovando, D., Free, C.M., Jensen, O.P., and Hilborn, R. 2022. A history and evaluation of catch-only stock assessment models. Fish and Fisheries, 23: 616-630.

Palomares, M.L.D., and Pauly, D. 2019. On the creeping increase of vessels' fishing power. Ecology and Society, 24:31.

Palomares, M.L.D., Froese, R., Derrick, B., Meeuwig, J.J., Nöel, S.-L., Tsui, G., Woroniak, J., Zeller, D., and Pauly, D. 2020. Fishery biomass trends of exploited fish populations in marine ecoregions, climatic zones and ocean basins. Estuarine, Coastal and Shelf Science, 243: 106896.

Patterson, K. 1992. Fisheries for small pelagic species: an empirical approach to management targets. Reviews in Fish Biology and Fisheries, 2: 321-338.

Pauly, D., Hilborn, R., and Branch, T.A. 2013. Fisheries: Does catch reflect abundance? Nature, 494, 303-306.

Pedersen, M.W., and Berg, C.W. 2017. A stochastic surplus production model in continuous time. Fish and Fisheries, 18: 226-243.

Pella, J.J., and Tomlinson, P.K. 1969. A generalized stock production model. Inter-American Tropical Tuna Commission Bulletin, 13: 419-496.

Perälä, T., and Kuparinen, A. 2017. Detection of Allee effects in marine fishes: analytical biases generated by data availability and model selection. Proceedings of the Royal Society B: Biological Sciences, 284: 20171284.

Perälä, T., Hutchings, J.A., and Kuparinen, A. 2022. Allee effects and the Allee-effect zone in northwest Atlantic cod. Biology Letters, 18: 20210439.

PEW (The Pew Charitable Trusts). 2016. Reference points: measuring success in fisheries management. [Accessed May 22, 2022].

PFMC (Pacific Fishery Management Council). 2014. Pacific coast groundfish fishery management plan for the California, Oregon and Washington Groundfish Fishery. Pacific Fishery Management Council, Portland, OR, USA.

Plagányi, É.E., Rademeyer, R.A., Butterworth, D.S., Cunningham, C.L., and Johnston, S.J. 2007. Making management procedures operational—innovations implemented in South Africa. ICES Journal of Marine Science, 64: 626-632.

Pons, M., Cope, J.M., and Kell, L.T. 2020. Comparing performance of catch-based and length-based stock assessment methods in data-limited fisheries. Canadian Journal of Fisheries and Aquatic Sciences, 77: 1026-1037.

Preece, A., Hillary, R., and Davies, C. 2011. Identification of candidate limit reference points for the key target species in the WCPFC. WCPFC-SC7-2011/MI-WP-03. 33pp.

Prince, J., and Hordyk, A. 2019. What to do when you have almost nothing: A simple quantitative prescription for managing extremely data-poor fisheries. Fish and Fisheries, 20: 224-238.

Prince, J D., Hordyk, A., Valencia, S.V., Loneragan, N., and Sainsbury, K. 2015. Revisiting the concept of Beverton-Holt life history invariants with the aim of informing data-poor fisheries assessment. ICES Journal of Marine Science, 72: 194-203.

Punt, A.E. 2008. Refocusing stock assessment in support of policy evaluation. In: Fisheries for Global Welfare and Environment, 5th World Fisheries Congress 2008. Eds: Tsukamoto, K., Kawamura, T., Takeuchi, T., Beard, T.D., and Kaiser, M.J. 139-152pp.

Punt, A.E., and Smith, A.D.M. 2001. The gospel of Maximum Sustainable Yield in fisheries management: birth, crucifixion and reincarnation. In: Conservation of Exploited Species (eds. J.D. Reynolds, G.M. Mace, K.H. Redford & J.G. Robinson), pp. 41-66. Cambridge University Press, Cambridge.

Punt, A.E., and Szuwalski, C. 2012. How well can FMSY and BMSY be estimated using empirical measures of surplus production? Fisheries Research, 134-136: 113-124.

Punt, A.E., Dorn, M.W., and Haltuch, M.A. 2008. Evaluation of threshold management strategies for groundfish off the U.S. West Coast. Fisheries Research, 94: 251-266.

Punt, A.E., Smith, A.D., Smith, D.C., Tuck, G.N., and Klaer, N.L. 2014. Selecting relative abundance proxies for B_{MSY} and B_{MEY}. ICES Journal of Marine Science, 71: 469-483.

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

Punt, A.E., Castillo-Jordan, C., Hamel, O.S., Cope, J.M., Maunder, M.N., and Ianelli, J.N. 2021. Consequences of error in natural mortality and its estimation in stock assessment models. Fisheries Research, 233: 105759.

Quinn, T.J., and Deriso, R.B. 1999. Quantitative fish dynamics. Oxford University Press, Oxford.

Rademeyer, R.A., Plagányi, É.E., and Butterworth, D.S. 2007. Tips and tricks in designing management procedures. ICES Journal of Marine Science, 64: 618-625.

Restrepo, V.R., Thompson, G.G., Mace, P.M., Gabriel, W.L., Low, L.L., MacCall, A.D., Methot, R.D., Powers, J.E., Taylor, B.L., Wade, P.R., and Witzig, J.F. 1998. Technical guidance on the use of precautionary approaches to implementing national standard 1 of the Magnuson-Stevens fishery conservation and management act. NOAA Technical Memorandum NMFS–F/SPO–31.

Reuchlin-Hugenholtz, E., Shackell, N.L., and Hutchings, J.A. 2016. Spatial reference points for groundfish. ICES Journal of Marine Science, 73(10): 2468-2478.

Ricard, D., Minto, D., Jensen, O.P., and Baum, J.K. 2012. Examining the knowledge base and status of commercially exploited marine species with the RAM Legacy Stock Assessment Database. Fish and Fisheries,13: 380-398.

Rice, J.C. and Rivard, D.S. 2007. The dual role of indicators in optimal fisheries management strategies. ICES Journal of Marine Science, 64: 775-778.

Ricker, W.E. 1954. Stock and recruitment. Journal of the Fisheries Research Board of Canada, 11: 559-623.

Rivard, D., and Rice, J., 2003. National workshop on reference points for gadoids. DFO Can. Sci. Advis. Sec. Proc. Ser. 2002/003. 16 pp.

Rosenberg, A.A., Fogarty, M.J., Cooper, A.B., Dickey-Collas, M., Fulton, E.A., Gutiérrez, N.L., et al. 2014. Developing new approaches to global stock status assessment and fishery production potential of the seas. FAO Fisheries and Aquaculture Circular No. 1086. Rome, FAO. 175 pp.

Rowe, G., and Wright, G. 2001. Expert opinions in forecasting: the role of the Delphi technique. In Principles of forecasting (pp. 125-144). Springer, Boston, MA.

Rudd, M.B., and Thorson, J.T. 2018. Accounting for variable recruitment and fishing mortality in length-based stock assessments for data-limited fisheries. Canadian Journal of Fisheries and Aquatic Sciences, 75: 1019-1035

Sainsbury, K. 2008. Best Practice Reference Points for Australian Fisheries. Australian Fisheries Management Authority Report R2001/0999.

Schaefer, M.B. 1954. Some aspects of the dynamics of the populations important to the management of the commercial marine fisheries. Inter-American Tropical Tuna Commission, 1: 27-56.

Schnute, J. 1985. A general theory for analysis of catch and effort data. Canadian Journal of Fisheries and Aquatic Sciences, 42: 414-429.

Schnute, J.T. and Richards, L.J. 2002. Surplus production models. In: Handbook of Fish Biology and Fisheries. P.J.B. Hart and J.D. Reynolds, eds. Blackwell Science Ltd, Oxford, UK. pp 105-126.

Serebryakov, V.P. 1991. Predicting year-class strength under uncertainties related to the survival in the early life history of some North Atlantic commercial fish. NAFO Science Council Studies, 16: 49-56.

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

Shelton, P.A., and Rice, J.C. 2002. Limits to overfishing: reference points in the context of the Canadian perspective on the precautionary approach. Can. Sci. Advis. Sec. Res. Doc. 2002/084. 29 pp.

Silvar-Viladomiu, P., Minto, C., Halouanu, G., Batts, L., Brophy, D., Lordan, C., and Reid, D.G. 2021. Moving reference point goalposts and implications for fisheries sustainability. Fish and Fisheries, 22: 1345-1358.

Simmonds, J., Hjorleifsson, E., and Millar, C. 2019. MSY: estimation of equilibrium reference points for fisheries. R package version 0.1.19. [Accessed May 22, 2022].

Sissenwine, M.P., and Shepherd, J.G. 1987. An alternative perspective on recruitment overfishing and biological reference points. Canadian Journal of Fisheries and Aquatic Sciences, 44: 913-918.

Smith, S.J., Bourdages, H., Choi, J., Dawe, E., Dunham, J.S., Gendron, L., et al. 2012. Technical guidelines for the provision of scientific advice on the precautionary approach for Canadian fish stocks: Section 7 – Invertebrate Species. DFO Can. Sci. Advis. Sec. Res. Doc. 2012/117. iv + 30 p.

Stenson, G.B., Hammill, M., Ferguson, S., Stewart, R., and Doniol-Valcroze, T. 2012. Applying the precautionary approach to marine mammal harvests in Canada. DFO Can. Sci. Advis. Sec. Res. Doc. 2012/107. ii + 15 p.

Szuwalski, C., and Punt, A.E. 2012. Identifying research priorities for management under uncertainty: the estimation ability of the stock assessment method used for eastern Bering Sea snow crab (*Chionoecetes opilio*). Fisheries Research, 134-136, 82-94.

Szuwalski C., Ianelli J. N., and Punt A. E. 2018. Reducing retrospective patterns in stock assessment and impacts on management performance. ICES Journal of Marine Science, 75: 596-609.

Then, A.Y., Hoenig, J.M., Hall, N.G., and 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: 82-92.

Thorson, J.T., Cope, J.M., Branch, T.A., and Jensen, O.P. 2012. Spawning biomass reference points for exploited marine fishes, incorporating taxonomic and body size information. Canadian Journal of Fisheries and Aquatic Sciences, 69: 1556-1568.

Thorson, J.T., Minto, C., Minte-Vera, C.V., Kleisner, K.M., Longo, C., and Jacobson, L. 2013. A new role for effort dynamics in the theory of harvested populations and data-poor stock assessment. Canadian Journal of Fisheries and Aquatic Sciences, 70: 1829-1844.

Thorson, J.T. 2020. Predicting recruitment density dependence and intrinsic growth rate for all fishes worldwide using a data-integrated life-history model. Fish and Fisheries, 21: 237-251.

Trijoulet, V., Berg, C.W., Miller, D.C.M., Nielsen, A., Rindorf, A., and Albertsen, C.M. 2022. Turning reference points inside out: comparing MSY reference points estimated inside and outside the assessment model. ICES Journal of Marine Science, 79: 1232-1244.

UN (United Nations). 1995. Agreement for the implementation of the provisions of the United Nations Convention on the Law of the Sea of 10 December, 1982 relating to the conservation and management of straddling fish stocks and highly migratory fish stocks. UN Rome, Italy.

UN. 2002. Report of the world summit on sustainable development. Technical report, A/CONF.199/20. United Nations Publication, Johannesburg, South Africa. 26 August - 4 September 2002.

van Deurs, M., Brooks, M.E., Lindengren, M., Henriksen, O., and Rindorf, A. 2021. Biomass limit reference points are sensitive to estimation method, time-series length and stock development. Fish and Fisheries, 22: 18-30.

Vasconcellos, M., and Cochrane, K. 2005. Overview of world status of data-limited fisheries: inferences from landings statistics. In: Kruse, G.H., Gallucci, V.F., Hay, D.E., Perry, R.I., Peterman, R.M., Shirley, T.C., Spencer, P.D., Wilson, B., Woodby, D. (Eds.), Fisheries Assessment and Management in Data-Limited Situations. Alaska Sea Grant College Program.

Vasilakopoulos, P., and Maravelias, C.D. 2016. Application of data-limited assessment methods on black anglerfish (*Lophius budegassa*; Spinola,1807) stocks in the Mediterranean Sea. Journal of Applied Ichthyology, 32: 277-287.

Vasilakopoulos, P., Jardim, E., Konrad, C., Rihan, D., Mannini, A., Pinto, C., Casey, J., et al. 2020. Selectivity metrics for fisheries management and advice. Fish and Fisheries, 21: 621-638.

Wakeford, R., Merino, G., Apostolaki, P., Skerritt, D., and Davies, T. 2019. Reference points, harvest control rules, and management strategy evaluation in tuna Regional Fisheries Management Organizations. WCPFC-SC15-2019/MI-WP-14 (Rev.01)

Walters, C.J. 1987. Nonstationarity of production relationships in exploited populations. Canadian Journal of Fisheries and Aquatic Sciences, 44(S2), s156–s165.

Walters, C., and Martell, S.J.D. 2002. Stock assessment needs for sustainable fisheries management. Bulletin of Marine Science, 70: 629-638.

Wiff, R., Flores, A., Neira, S. and Caneco, B. 2018. Estimating steepness of the stock-recruitment relationship in Chilean fish stocks using meta-analysis. Fisheries Research, 200: 61-67.

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.

Worm B., Barbier, E.B., Beaumont, N., Duffy, J.E., Folke, C., Halpern, B.S. et al. 2006. Impacts of biodiversity loss on ocean ecosystem services. Science, 314: 787-790.

Zhou, S., Yin, S., Thorson, J.T., Smith, A.D., and Fuller, M. 2012. Linking fishing mortality reference points to life history traits: an empirical study. Canadian Journal of Fisheries and Aquatic Sciences, 69: 1292-1301.

Zhou, S., Punt, A.E., Ye, Y., Ellis, N., Dichmont, C.M., Haddon, M., Smith, D.C., and Smith, A.D.M. 2017. Estimating stock depletion level from patterns of catch history. Fish and Fisheries, 18: 742-751.

Zhou, S., Punt, A.E., Smith, A.D.M., Ye, Y., Haddon, M., Dichmont, C.M., and Smith, D.C. 2018. An optimized catch-only assessment method for data poor fisheries. ICES Journal of Marine Science, 75: 964-976.

APPENDIX A: DATA-RICH ESTIMATION METHODS FOR B_0 AND B_{MSY}

Here, we define data-rich methods to estimate B_0 and B_{MSY} as those used in analytical assessments (i.e., supported by age-structured, size-structured, or surplus production models). Estimates of stock parameters (e.g., *h* and *M*) and reference points can still be highly uncertain in data-rich scenarios. Stock-recruit estimates are generally highly variable and the functional form of the SRR is usually unknown. Estimates of B_{MSY} depend on model assumptions and proxies for B_{MSY} may be more reliably estimated. Various thresholds for recruitment overfishing have been defined based on replacement *F*, change points estimated from segmented regressions, and generic rules of thumb (e.g., 0.2 B_0 or biomass at 50% of maximum model estimated recruitment from a SRR). Potential solutions to capturing uncertainty in the SRR in estimates of stock status include fitting multiple assessment models with different SRRs or capturing different hypotheses about the SRR in alternative operating models. Estimates of stock status from multiple models can be combined using ensemble (model averaging) methods.

A.1. AGE- OR SIZE-STRUCTURED MODELS WITH A STOCK RECRUITMENT RELATIONSHIP

A.1.1. Stock Recruitment Relationships

Stocks that use age- or size-structured analytical models generally assume a SRR and describe the population and fishery exploitation pattern based on age- or size-structure. This section addresses SRR for iteroparous species but semelparous species also use SRRs for which reference points can be estimated directly from the SRR parameter estimates (see Section 2.8). A common SRR is the BH model (Beverton and Holt 1957):

$$R(B) = \frac{\alpha B}{1+\beta B} \tag{A.1}$$

where recruitment (*R*) is modelled as a function of two parameters α (the slope near the origin, or maximum survival rate) and β (which determines the degree of density-dependence), and biomass (*B*, usually SSB) for which recruitment increases to an asymptotic value of α/β.

Another common SRR is the Ricker model (Ricker 1954):

$$R(B) = \alpha B e^{-\beta B} \tag{A.2}$$

also a function of α and β. The α has the same interpretation as the BH model, while the β parameter is also a density-independent parameter that determines the degree of density dependence, including “over-compensation”, where *R* can decline at high biomass (Figure A.1). Alternative parameterizations of (A.1) and (A.2) are also commonly used (e.g., Hilborn and Walters 1992; Mangel et al. 2010), including parameterizations based on *h*. For example, a version of the BH model using the *h* parameterization is (Dorn 2002; Miller and Brooks 2021):

$$R(B) = \frac{4R_0 hB}{(1-h)R_0\varphi_0+(5h-1)B} \tag{A.3}$$

where φ_0 is the unfished (spawning stock) *B* per recruit (where $\varphi_0 = B_0/R_0$). A version of the Ricker model using the *h* parameterization is (Miller and Brooks 2021):

$$R(B) = \frac{B}{R_0}(5h)^{\frac{5}{4}\left(1-\frac{B}{R_0\varphi_0}\right)} \tag{A.4}$$

The parameterization of the SRR in terms of *h* is convenient for practitioners due to the interpretation of *h* in terms of the resilience of the stock (greater resilience corresponds to higher *h*), where *h* is a standardized measure of productivity that appears to be comparable among stocks (Dorn 2002; Miller and Brooks, 2021). The *h* parameterization of the BH SRR implicitly assumes that the productivity of recruited individuals (through the φ_0 term of Eqn. A.3 and A.4) is constant, in addition to the assumed constant pre-recruit mortality rate. With estimates of *h* and R_0, one can derive B_0 (from Eqn. A.3 or A.4) and specify the SRR. Estimates of *h* can be supported from results obtained from meta-analyses (e.g., Myers et al. 2002; Dorn 2002; Michielesens and McAllister 2004; Forrest et al. 2010; Sherzer and Conn 2012; FishBase, Froese and Pauly 2022) or *h* can be defined explicitly or as a range of plausible values in a Bayesian approach (Mangel et al. 2010). Commonly, one of these approaches is used to develop a prior probability distribution within a Bayesian stock assessment model.

In some cases, SRR parameters may be time-varying (van Deurs et al. 2021), making estimation of reference points challenging. Miller and Brooks (2021) note that many applications of SRRs violate the important assumptions underlying the derivation of the SRR (e.g., constant φ_0). When the components of φ_0 (i.e., *M*-at age, maturity-at-age, and weight-at-age) change over time then *h* implicitly also changes over time. In this situation the α, β parameterization of the SRR could be one approach to avoid misinterpretation of the SRR (Miller and Brooks 2021) and resulting bias in the estimates of MSY reference points.

The variability in the SRR data, uncertainty in the functional form of the SRR, and/or uncertainty in the SRR parameter estimates may be so large that candidate LRPs that rely on the SRR parameter estimates may be unreliable (Myers et al. 1994; van Deurs et al. 2021). Reliability of B_{MSY} estimates depend strongly on model parameters, especially *h* and *M*. When the assumptions concerning *h* of the SRR are unsupported, alternative metrics and proxies for B_{MSY} (e.g., proportions of B_0 or equilibrium biomass at $F_{\text{X\%SPR}}$; Section 2.2) may be more credible than direct estimates of B_{MSY} (MF 2011). Similarly, when *h* is fixed (i.e., not estimated in an assessment model), B_{MSY} should not be used as the basis for reference points (Punt 2008). Trijoulet et al. (2021) suggested that MSY-based reference points should not be used if the SRR is poorly estimated or the MSY reference points are not estimated with confidence. This is echoed by the Western and Central Pacific Fisheries Commission (WCPFC) guidance for defining LRPs based on the reliability of *h* in the SRR (Preece et al. 2011). The WCPFC recommends: using a proportion of B_{MSY} if a reliable estimate of *h* is available; 0.2 B_0, 0.2 B_{unfished}, or a proportion of the equilibrium biomass based on $F_{\text{\%SPR}}$ if *h* is not well estimated or is unknown; and 0.2 B_0 or 0.2 B_{unfished} if key biological and fishery parameters are not well estimated or known (Preece et al. 2011).

High variability in recruitment is typical of short-lived, fast-growing, and early maturing small pelagic fish (van Deurs et al. 2021). Extrinsic factors may be more important drivers of recruitment than SSB for some stocks. For example, Brosset et al. (2019) found that physical and biological environmental variability was a significant predictor of recruitment of spring spawning herring (*Clupea harengus*) in the Gulf of St. Lawrence, while SSB was not a significant predictor of recruitment. Short-lived stocks with volatile dynamics may pose challenges for LRP estimation due to changes in the SRR over time (van Deurs et al. 2021).

High variability in recruitment and a lack of stock-recruit data at low biomass can mask the underlying SRR. When there is no SRR apparent in the stock-recruit data, the assumption of *h* approaching 1 implies productivity approaching infinity which is biologically unreasonable, and also risk-prone in terms of setting reference points (Mangel et al. 2013). If a high *h* is assumed, B_{MSY} estimates can be extremely low (Myers et al. 1994) and candidate LRPs based on proxies may need to be considered (see Sections B.1 and 2.4).

An alternative approach to estimating stock status that addresses the uncertainty in the SRR is to combine estimates from multiple models with different structural assumptions using model ensemble approaches (see Section 4.2).

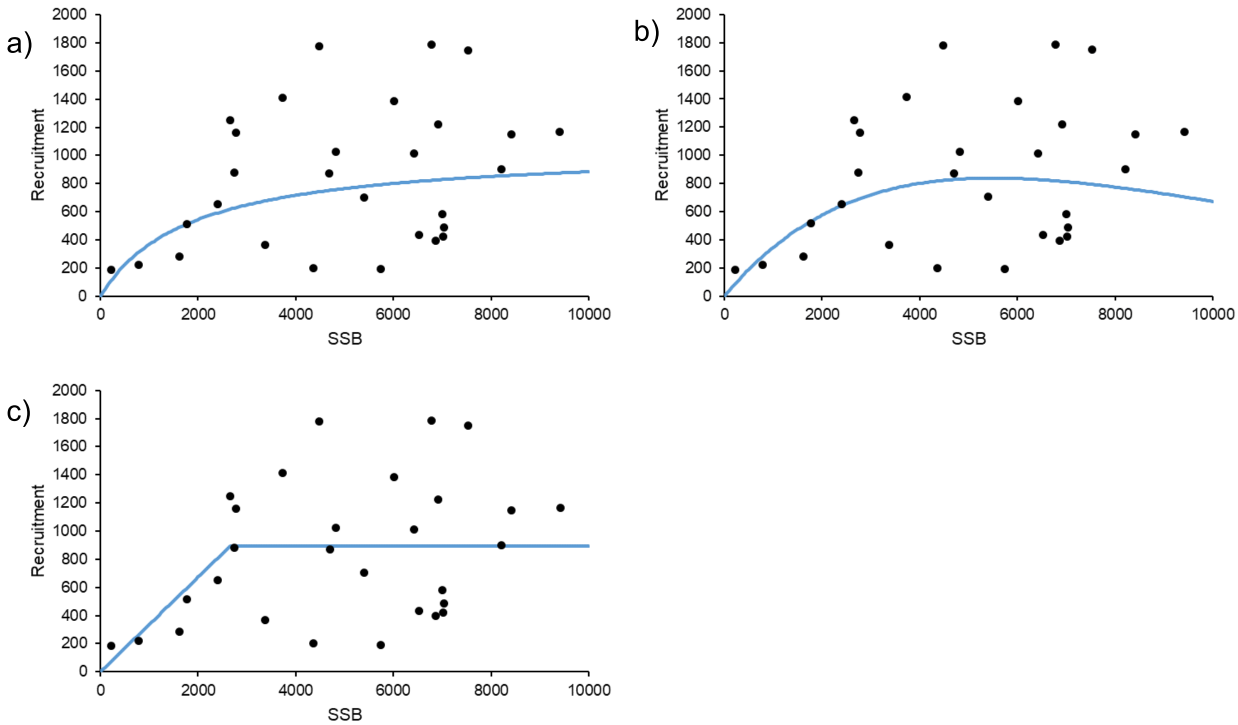

Figure A.1. Example dataset of stock (SSB) and recruit (recruitment) pairs fit to different functional forms of the stock recruitment relationship a) Beverton-Holt, b) Ricker, and c) hockey stick (segmented regression)

A.1.2. Natural Mortality Rate

M is considered one of the most important parameters in stock assessment but is difficult to estimate (Hamel 2015; Then et al. 2015; Punt et al. 2021). Uncertainty in *M*, is influenced by time varying processes such predation, food availability, disease, and environmental effects (Punt et al. 2021). An additional challenge in estimating *M* is that it is confounded with *h* and the growth-rate, and it also declines in selectivity with increasing age (Punt et al. 2021). *M* is a function of age and sex, but commonly assumed to be a constant in assessment models. *M* may be fixed, or estimated within an assessment model for which priors on *M* may be assumed. Estimates of *M* influence estimates of B_0 and MSY reference points. In general a higher *M* results in a higher estimate of B_0, a higher F_{MSY}, a lower estimate of B_{MSY}, and a lower ratio of B_{MSY}/B_0 (He et al. 2011; Szuwalski and Punt 2012; Punt et al. 2021). Although the magnitude of the reference points are influenced by *M*, estimates of B_{MSY} depend strongly on *h* and are not impacted substantially by *M* when *h* is correctly specified (Punt et al. 2021). He et al. (2011) conducted a simulation study using an age-structured model and found that misspecification of *M* influenced B_0 estimates but had a negligible effect on depletion estimates (i.e., changes in *M* scale both current biomass and B_0). However, when *M* is time-varying, bias in *M* can influence depletion estimates if terminal year *M* differs from *M* assumed over the period used to define B_0 (Johnson et al. 2015).

A.1.3. Selectivity

Fishery selectivity-at-age relative to maturity-at-age influences estimates of MSY-based reference points (Myers and Mertz 1998; Punt et al. 2014). Increased yields and lower probabilities of stock collapse have been demonstrated by adjusting the selectivity-at-age or selectivity-at-size, usually to reduce the harvest of juveniles (e.g., Myers and Mertz 1998, Froese et al. 2016; Vasilakopoulos et al. 2016; Prince and Hordyk 2019). The mechanism through which increasing the age- or size-at-first-capture promotes sustainability is by preventing growth overfishing and recruitment overfishing (Vasilakopoulos 2020). There is a trade-off between fishing mortality rate and selectivity, such that a higher *F* can be sustained when catching larger fish but catching too many juvenile fish can lead to stock depletion at moderate levels of *F* (Prince and Hordyk 2019; Vasilakopoulos et al. 2016). Therefore, a selectivity that targets larger fish will lead to a higher F_{MSY} and therefore a lower B_{MSY}. Changes to selectivity do not influence estimates of B_0.

A.2. SURPLUS PRODUCTION MODELS

In the absence of size- or age-structured data, surplus production models implicitly capture the combined effects of recruitment, growth, and *M* in a single parameter *r* (the intrinsic rate of population growth). In these models, the exploitation rate is simply the ratio of catch to stock biomass so the stock biomass is assumed to be the exploitable biomass. Surplus production models rely on the assumption that the predicted surplus production of a population depends on population size. Production models require at minimum a times series of catch-per-unit-effort (CPUE) and total catch and work best when there is considerable contrast in effort, biomass, and catch rate (Hilborn and Walters 1992; ICES 2018). These models can be criticized for being overly simplistic since they do not account for size-structure, varying recruitment, or changes in selectivity, which can result in lags in population responses to perturbations such as fishing (DFO 2016); however, they may be the best available approach when only size-aggregated data are available and SRRs cannot be estimated. MSY reference points can be estimated directly from surplus production models. The most common surplus production model is based on the Schaefer (1954) production model:

$$B_{t+1} = B_t + rB_t\left(1 - \frac{B_t}{K}\right) - C_t \tag{A.5}$$

where B_t in this case represents total biomass at time *t*, *r* was defined above, *K* represents the population carrying capacity, and C_t is the catch removed in year *t*. This version of the surplus production model assumes a symmetric production vs. biomass curve such that $B_{\text{MSY}} = 0.5\ K$. More flexible production models exist that deviate from the symmetric production vs. biomass curve. Another common model is based on the Fox (1970) production model:

$$B_{t+1} = rB_t ln\,(K)\left(1 - \frac{ln\,(B_t)}{ln\,(K)}\right) - C_t \tag{A.6}$$

which has a production vs. biomass curve that is skewed to the right such that $B_{MSY} \sim 0.37\ K$ ($0.4\ B_{\text{MSY}} = 0.144\ B_0$). The Pella-Tomlinson production model (Pella and Tomlinson, 1969) is more flexible with an additional shape parameter that is equivalent to the Schaefer production model when $m = 2$ and reduces to the Fox model as *m* approaches 1 (Quinn and Deriso 1999):

$$B_{t+1} = B_t + \frac{r}{m-1}B_t\left(1 - \left(\frac{B_t}{K}\right)^{m-1}\right) - C_t \tag{A.7}$$

B_{MSY} can be calculated directly from the *m* parameter (Table A.1). The default DFO PA Policy guidance of $0.4\ B_{\text{MSY}}$ as an LRP can result in low depletion for the Fox and Pella-Tomlinson models (DFO 2016). An LRP of $0.2\ K$ was suggested as an LRP for these models by DFO (2016).

Table A.1. Estimates of B_{MSY} and F_{MSY} from biomass dynamics models.

Model	B_{MSY}	F_{MSY}
Schaefer	$0.5\,K$	$0.5\,r$
Fox	$e^{-1}K$	r
Pella-Tomlinson	$m^{-\left(\frac{1}{m-1}\right)}K$	$\frac{r}{m-1}\left(1-\frac{1}{m}\right)$

Notes: K = carrying capacity (B_0); r = intrinsic rate of population growth rate; m = Pella-Tomlinson shape parameter; e = Euler's number; MSY is estimated as the product of B_{MSY} and F_{MSY}.

Criticisms of surplus production models include the possible problem that they do not account for reductions in the rate of recruitment at low biomass (Froese et al. 2017) and that they assume that the rate of change of biomass over time increases as biomass approaches zero (Schnute and Richards 2002). A solution to this shortcoming is to add a SRR to the production model. There are multiple approaches to do this (e.g., McAllister et al. 2001; Schnute and Richards 2002, Mangel et al. 2013). A simple approach proposed by Froese et al. (2016) assumes constant recruitment above a proportion (P_{lim}) of K and that recruitment declines linearly below P_{lim} to zero, consistent with the hockey stick SRR (ICES 2021a; Figure A.1). This approach results in the following model (Froese et al. 2016):

$$B_{t+1} = B_t + \frac{r}{m-1}\left(\frac{B_t}{P_{lim}K}\right)B_t\left(1-\left(\frac{B_t}{K}\right)^{m-1}\right) - C_t \qquad \text{if } \frac{B_t}{K} < P_{lim} \qquad \text{(A.8)}$$

where P_{lim} in the range of 0.2 – 0.25 represents a threshold for recruitment overfishing (i.e., recruitment declines linearly to zero below 0.2 – 0.25 B_0*)*.

The stochastic surplus production model in continuous-time (SPiCT; Pedersen and Berg, 2017) is a more advanced model based on the Pella-Tomlinson formulation but allows for observation error in catch and includes one or more abundance indices (e.g., CPUE, biomass, survey indices) to estimate stock parameters such as depletion, biomass, and *F*. The continuous-time formulation of the model can accommodate arbitrary and irregular data sampling (e.g., when catch and biomass indices do not match temporally). The surplus production models in SPiCT use biomass indices to estimate *K*, which helps scale the trends in *F* and biomass (instead of relying on depletion priors like catch-only methods; Appendix C). The *K* parameter cannot be well estimated when the index time series does not extend back to the beginning of the fishery (Punt and Szuwalski, 2012) or does not have sufficient contrast in biomasses over the time series (Ono et al. 2012). Priors for depletion at the beginning of the time series are important when the catch and index time series begin after a period of high exploitation (Bouch et al. 2021). SPiCT models provide estimates of MSY reference points and include estimates of uncertainty (95% confidence intervals). Bouch et al. (2021) compared SPiCT evaluations to ICES stock status determination from 17 assessed stocks from the Celtic, Irish, and North Seas and found that SPiCT predicted lower relative *F* (15 of 17 stocks) and higher relative biomass (15 of 17 stocks) when compared to their respective ICES assessment results, which suggests an increased risk of generating overly optimistic results. The SPiCT modelling package (Pedersen and Berg, 2017) includes diagnostics to evaluate model assumptions. Additionally, SPiCT is included as a suggested resource in the guidance for ICES Category 3 and 4 stocks (ICES 2018).

An intermediate between age-structured models and biomass-dynamic models is the delay-difference model (Deriso 1980; Schnute 1985; Hilborn and Walters 1992). The delay-difference model is essentially a collapsed age-structured model that tracks the effects of recruitment, survival, and growth on biomass, without requiring a fully age-structured framework, and can perform well, as long as its major assumptions are met, which include growth in mean body weight follows the Ford-Walford equation; fishery selectivity-at-age and maturity-at-age are knife-edged (i.e., fish mature at the same age, a_k, they become vulnerable to the fishery, and all fish aged a_k and older are equally vulnerable to fishing gear); and constant *M*-at-age (Hilborn and Walters, 1992). Reference points in the delay-difference model are calculated using the same equations as for age-structured models, with the advantage that the models can be fit without size or age composition data (e.g., Forrest et al., 2020). However, it is advantageous if there are enough size and age data to estimate growth parameters and to have an understanding of the likely age at which fish mature and become vulnerable to the fishery. Estimates of MSY-based reference points made using delay-difference models are subject to the same uncertainties described for age-structured models (e.g., uninformative SRR) and may also be biased if the assumptions of the delay-difference model are violated. For example, Forrest et al. (2018) used a delay difference model to estimate the biomass of Pacific cod (*Gadus macrocephalus*) stocks, but used reference points based on historical estimates of biomass and fishing mortality.

APPENDIX B: DATA-LIMITED ESTIMATION METHODS FOR B_0 AND B_{MSY}

Here, we define data-limited methods to estimate stock status as those that lack data or the quality of data typical of the data-rich methods in Appendix A (i.e., age- or size-structured data or the index of abundance used in biomass dynamics models). These include using empirical or theoretical proxies for B_0 and B_{MSY}, historical biomass levels, catch-only methods, and evaluating the performance of harvest strategies using closed-loop simulations. Many Canadian stocks that lack reference points are data-limited. Empirical and theoretical proxies for B_0 and B_{MSY} may be estimated in some cases. Catch-only methods can provide estimates of stock status relative to B_0 and B_{MSY}; however, they have generally been used to estimate global trends in stock status and have been shown not to be a reliable method of assessing individual stocks. A closed-loop simulation approach is a robust alternative to using empirical proxies (that may not capture uncertainty in the estimate of stock status) or methods that rely on strong model assumptions (that may not be valid) or priors (that may not be reliably determined).

B.1. PROXIES FOR B_0 AND B_{MSY}

B.1.1. Theoretical Proxies for B_{MSY}

The *F* that maximizes yield-per-recruit (YPR) is F_{max} and it has been used as a basis for reference points. However, F_{max} represents a threshold for growth overfishing (where fish are harvested before they reach a size that enables the population to reach biomass levels needed to achieve MSY) and, as such, may exceed a threshold for recruitment overfishing (Shelton and Rice 2002). F_{max} is always greater than or equal to F_{MSY}, which can lead to overfishing and result in excessive depletion (Sainsbury 2008) and is therefore not recommended as a proxy for F_{MSY}. A more conservative alternative to F_{max} is $F_{0.1}$ (proposed by Gulland and Boerema 1973) which is the level of *F* found where the slope of the YPR vs. *F* curve is 10% of the slope of the curve at its origin. DFO's PA Policy suggests the equilibrium biomass at $F_{0.1}$ as a possible proxy for B_{MSY} (DFO 2009), but the choice of 10% is arbitrary, sensitive to assumptions about growth and *M*. $F_{0.1}$ can overestimate F_{MSY} in some cases (Mace and Sissenwine 1993, Myers et al. 1995b). Mace (2001) recommended against using $F_{0.1}$ without understanding how well it relates to F_{MSY}.

B.1.2. B_{MSY} Proxies based on Natural Mortality Rate (M)

A general rule of thumb has been to assume that F_{MSY} is approximately equal to M and therefore B_{MSY} is approximately equal to the equilibrium biomass at $F = M$ (e.g., Francis 1974, Mace 1994; Froese et al. 2016). However, it has been noted that such proxies are often considered in largely data-limited scenarios where it may not be possible to reliably estimate M either (Gabriel and Mace 1999). Zhou et al. (2012) conducted a meta-analysis of 245 fish species to evaluate the relationships between F_{MSY} (and F_{MSY} proxies) and life-history parameters, including M. They concluded that M was the most important life-history parameter influencing F reference points, and that the best model for teleosts was $F_{MSY} = 0.87M$ (standard deviation = 0.05). While this represents an average result from a meta-analysis, F_{MSY} has been shown to be closer to $0.5M$ for some stocks (Walters and Martell 2002), especially pelagic species (Patterson 1992).

B.1.3. Historical, Empirical, or Other Proxies for B_0 and B_{MSY}

Candidate historical or empirical LRPs can be defined based on:

- a B_{MSY} proxy, defined by
  - the mean or median value of an indicator over a historical time period when the indicator is high and catches are high; or
  - the mean or median value of an indicator over a productive period.
- a B_0 proxy, defined by
  - the mean/median (or maximum) value of the indicator over a historical time period reflecting the beginning of exploitation.
- A historical low biomass state:
  - from which the stock recovered to above average levels;
  - or that is agreed by managers and resource users to be an undesirable state to avoid.

It may be reasonable to select a historical time period to approximate either B_0 or B_{MSY} using either estimated abundance or biomass, or an empirical stock indicator (e.g., survey index or CPUE) that is assumed to be proportional to biomass or abundance. An understanding of historical fishing pressure is needed to properly interpret the historical proxies for either B_0 or B_{MSY} (NAFO 2004). Understanding the relationship between an empirical indicator and biomass (or abundance) and accounting for environmental variability are important when considering potential indicators of stock status. For example, Murray and Seed (2010) found that CPUE could not be used as a proxy for green crab (*Carcinus maenas*) abundance due to the influence of temperature on crab activity levels and catches. Standardization of an indicator (by season or temperature) may be necessary to ensure the indicator is reflective of abundance (e.g., Cook et al. 2020).

The maximum population size (estimated or inferred) can be used as a historical proxy for B_0 or K (e.g., marine mammals; Stenson et al. 2012, DFO 2016), particularly if the indicator starts at the beginning of the exploitation time series. A provisional estimate for B_{MSY} is 50% of the maximum population size as suggested by the PA Policy (DFO 2009). If there are no data on recruitment, some function of CPUE might be used, but only if information on resource conditions prior to or shortly after the onset of fishing are available (Gabriel and Mace 1999, Restrepo et al. 1998).

If the stock already has some history of exploitation and has declined from B_0 to a stable biomass assumed to be B_{MSY}, then the highest observed indicator value may instead serve as a proxy for B_{MSY} (NAFO 2004). Therefore, the use of the highest observed indicator value as a proxy for B_0 or B_{MSY} relies on expert judgement on the interpretation of the ratio of B/B_0 at the beginning of the index time series.

As an alternative to using a maximum value as a proxy for B_{MSY}, the mean or median value of an indicator over a productive period (DFO 2009) may serve as a proxy for B_{MSY} (DFO 2016), given that B_{MSY} is attained when fishing at F_{MSY} on *average*, under equilibrium conditions. The choice of productive period can be subjective (Smith et al. 2012) but should be from a period when both indicators and catches are high (MF 2011). If both biomass estimates and catch data are available, estimates of surplus production can be investigated for evidence of MSY (DFO 2016). If a reference period is used to approximate MSY from fishery-dependent indices such as CPUE, catches or landings indicators, there should be no evidence that abundance was declining during that time (Gabriel and Mace 1999). That is, both CPUE and catches should have been relatively high and CPUE is considered reasonably proportional to stock size (MF 2011), suggesting biomass was relatively high and recruitment stable (DFO 2008). Regardless of the approach chosen, expert judgement is needed to support the choice of LRP.

LRPs can also be based on estimated historical biomass states without consideration of their relationship to B_0 or B_{MSY}. For example, a historical LRP for Pacific rock sole (*Lepidopsetta* spp.) was set at the minimum biomass for the period 1996-2005, when the stock was agreed to be low (DFO 2014; Holt et al. 2016). The LRP for Area 3CD Pacific cod (Forrest et al. 2020; DFO 2019b) is based on the estimate of the lowest historical biomass from which the stock recovered to above average levels. The LRP for Area 5ABCD Pacific cod is based upon an agreed historical low biomass state to be avoided, where the year of this state (2000) was agreed at the regional peer review meeting (DFO 2019b).

B.2. CATCH-ONLY APPROACHES

Many stock assessment methods have been developed for data-limited stocks that rely only on catch data. These methods can produce estimates of biomass relative to B_{MSY} or B_0. There is, however, controversy over the use of catch-only methods for estimating stock status (e.g., Branch et al. 2011, Pauly et al. 2013). Catch-only methods assume that trends in catch are indicative of trends in abundance or biomass. In developing fisheries, the catch is assumed to be initially low, increase over time in heavily exploited fisheries, decrease in overexploited fisheries, and decline when the fishery collapses (Branch et al. 2011). This assumption has been debated (e.g., Worm et al. 2006, Branch et al. 2011) since there are many other factors that may influence catch rates (e.g., management regulations, changes in market demand and fuel prices, natural disasters; Branch et al. 2011; Kleisner et al. 2013). Furthermore, data-limited stocks sometimes only have data on landings and assumptions must be made regarding the magnitude of discards (Rosenberg et al. 2014).

A recent review of 11 data-limited methods for estimating stock status evaluated their performance on data-rich stocks as well as simulated fisheries data (Free et al. 2020). The authors found that the catch-only methods they evaluated produced biased and imprecise estimates of relative depletion, especially for stocks that were lightly exploited. Catch-only methods have been developed primarily to assess the global status of unassessed stocks (e.g., Costello et al. 2012; Kleisner et al. 2013; Palomares et al. 2020). Given current interest in the application of catch-only methods for estimating stock status, a brief review of the catch-only methods described by Free et al. (2020) is provided in Appendix C. However, these methods have not proven to be a consistently reliable method of assessing individual stocks (Carruthers et al. 2014; Free et al. 2020; Sharma et al., 2021; Ovando et al. 2022). Therefore, other approaches should be preferred when possible (e.g., closed-loop simulation approaches; Carruthers and Hordyk 2019, see Section 4.4.2), efforts to collect data that enable the use of other assessment methods should be prioritized (Ovando et al. 2021), and catch-only methods should only be used after careful consideration of their assumptions and chosen priors.

APPENDIX C: REVIEW OF CATCH-ONLY METHODS

C.1. GRAPHICAL APPROACHES (STOCK STATUS PLOTS)

Catch-based Stock Status Plots (SSPs) are used to define the status of a fishery relative to the maximum catch in the time series. Various criteria and status categories have been defined by different authors (e.g., Froese and Kesner-Reyes (2002), Kleisner and Pauly (2011), Kleisner et al. (2013)). The categories defined by Kleisner et al. (2013) are presented in Table C.1. The SSP methods have been evaluated using simulation testing and were demonstrated to be poor and inherently pessimistic predictors of stock status (Branch et al., 2011; Carruthers et al., 2012). Modifications to the methods to overcome some of the pitfalls have been proposed (e.g., smoothing the catch time series; Anderson et al. 2012), but there remain limitations of the methods because the classification of stock status is based on assumed relationships between abundance, fishing effort, and catch over time. It is assumed that abundance first decreases as catch and effort increase in the developing phase. As the fishery becomes fully exploited and moves to over-exploited, the relationship between catch and abundance breaks down until high effort leads to lower abundance and catches eventually decline (Free et al. 2020).

The general application of the SSP methods has been to assign stock statuses for many stocks to infer global trends (e.g., Anderson et al. 2012, Kleisner et al. 2013) and the methods were not intended to be used to assess single stocks for fisheries management (Froese et al. 2022). While Kleisner et al. (2013) acknowledge that it can be difficult to attribute a change in catch to changes in biomass alone, they note that there would be few instances where a large decline in catch (e.g., changing from exploited to over-explored or collapse categories) would not be related to a decline in biomass.

The status categories from the SSP method (e.g., Table C.1) are defined based on the maximum catch. A stock status indicator could be defined to be consistent with the intent of an LRP (e.g., a proportion of maximum catch). The advantage of the method is that it is simple, however, the disadvantage is the dependence on a relationship between abundance, fishing effort, and catch over time. This method is not recommended unless these assumptions can be supported by empirical evidence or expert judgement.

Table C.1. Categories used by Kleisner et al. (2013) to define the status of fisheries based on catch time series.

Status	Criterion
Developing	Years where catch < year of maximum catch AND catch is ≤ 50% of maximum catch OR year of maximum catch = final year of catch
Exploited	Years where catch > 50% of maximum catch
Over-exploited	Years where catch > year of maximum catch AND catch is 10-50% of maximum catch
Collapsed	Years where catch > year of maximum catch AND catch < 10% of maximum catch

Status	Criterion
Rebuilding	Years where catch > year of minimum catch post maximum catch AND the post-maximum minimum catch < 10% of maximum catch AND catch is 10-50% of maximum catch

C.2. EMPIRICAL APPROACHES

Empirical models predict stock status (B/B_{MSY} or relative depletion) from catch and other information (e.g., life-history parameters, location, or fishery characteristics) using models that are fit using data for data-rich stocks (Free et al. 2020). The empirical approaches derive relationships between stock status and catch data to inform stock status from catch data of data-limited stocks in an attempt to learn from experience. Methods are typically cross-validated by testing them on a set of data-rich stocks.

The “only reliable catch stock” (ORCS) working group method estimates stock status (under, fully, or overexploited) based on the mean of 14 categorical predictor variables for the stock and fishery (Berkson et al. 2011). The method was further refined by Free et al. (2017) by using a boosted classification tree model that was trained on data-rich stocks in the RAM Legacy Database (Ricard et al. 2012), using 12 of the 14 original predictors. The refined ORCS method correctly classified 74% of stocks in the training dataset and 74% of stocks in a test dataset and out-performed other methods evaluated by Free et al. (2017); however, the classification rate was lower for overexploited stocks (62% and 50%, respectively). The refined ORCS method can be applied with missing data for any of the 12 predictor variables and the most important predictor variables were value of the fishery ($/lb), status of assessed stocks in the fishery, targeting intensity, and discard rate and occurrence in the catch.

An approach to estimating stock status using a panel regression model (PRM) was developed by Costello et al. (2012). PRMs were defined to predict B/B_{MSY} for data-limited stocks using the catch time series and predictor variables for life-history traits. The models were defined using data from a set of assessed fisheries (from the RAM Legacy Database; Ricard et al. 2012) then used to estimate B/B_{MSY} for data-limited fisheries. The method assumes that the status of a stock is a function of its life-history traits and harvest history and that the way these parameters influence stock status is consistent across stocks with similar characteristics (Costello et al. 2012). A modified PRM was defined by Rosenberg et al. (2014) and is a simplified version of the PRM using only the catch time series and a categorical predictor variable for life-history category (demersal, small pelagic, and large pelagic). The modified PRM can be applied to more stocks given it doesn’t require detailed life-history information.

A catch-only boosted regression tree model was developed by Zhou et al. (2017). It uses predictor variables based on catch history to estimate relative depletion and was trained on data-rich stocks in the RAM Legacy Database. Overall, the model out-performed the modified PRM and catch-based classification (Kleisner et al. 2013) methods, but the overall classification was driven by results for fully exploited stocks. Classification for developing, over-exploited, and collapsed stocks were low (8%, 32%, and 52%, respectively), compared to 72% for fully exploited stocks. Zhou et al. (2017) describe the application of the catch-only boosted regression tree model as a method to estimate a final year depletion prior to be used in other catch-only methods, and not as a method for estimating stock status.

As with the SSP method of estimating a stock status, the empirical approaches described here were intended to estimate stock status to infer global trends and not to assess individual stocks. A limitation of these methods is that they assume empirical relationships between catch history and observed stock status. They rely on the assumption that these relationships exist and are consistent between stocks with formal assessments (heavily managed stocks) and those that are unassessed. The empirical methods are therefore not recommended for estimating indicators of stock status (e.g., B/B_{MSY} or B/B_0). Although the empirical approaches may not be reliable for estimating stock status for individual stocks, they may be useful for estimating priors for relative depletion for other methods (Free et al. 2020) or inclusion into ensemble methods (e.g., Anderson et al. 2017; see below).

C.3. MECHANISTIC APPROACHES

Mechanistic approaches are those that rely on an underlying population dynamics model. These approaches fit a population dynamics model or a coupled model of population and fishing effort dynamics if effort data are also available (Free et al. 2020). Assumptions are required for fishing effort and priors are generally set for initial and final year depletion, intrinsic rate of population growth (r), and carrying capacity (K). A range of reference points can be estimated including B/B_{MSY} and F/F_{MSY}.

C.3.1. Catch-MSY

The catch-MSY family of methods (Catch-MSY, Martell and Froese 2013; CMSY, Froese et al. 2017; and CMSY++, Froese et al. 2022) are based on the Schaefer production model (Equation A.5; Appendix A). Initial estimates of r and K are required (e.g., r based on life-history traits from fishbase.org, Froese and Pauly 2022; or Fishlife, Thorson 2020) and the time series of biomass and catch can be projected with $F_{MSY} = r/2$ and $B_{MSY} = K/2$ (Froese et al. 2022) using stock reduction analyses that reconstruct historical biomass and F by estimating biomass trajectories that could have produced the historical catch time series. The catch-MSY method generates thousands of biomass trajectories from “viable” r-K pairs (note that r and K are negatively correlated) that did not lead to a population outside of the expected depletion range (Froese et al. 2022).

Catch-MSY was updated to CMSY (Froese et al. 2017) with a modification to the method of identifying the most probable r-K pairs. The 75th percentile of viable r values was selected (instead of the center of the viable r values) and a central value with the confidence limits of the selected r values was used (Froese et al. 2022). The modification of the method resulted in improved estimates of B/B_{MSY} and F/F_{MSY}. CMSY was also revised to account for reduced recruitment at low biomass (i.e., $B/K < 0.25$), by including a stock recruitment relationship (Equation A.8, Appendix A). CMSY was updated to CMSY++ (Froese et al. 2022) with some improvements to the default priors, the method to estimate r-K pairs (now sampled from a multivariate lognormal distribution that accounts for the negative correlation between r and K), and an option to consider increases in CPUE over time not related to increases in biomass (i.e., technological creep, Palomares and Pauly, 2019).

Ovando et al. (2021) found that estimates of the status of stocks based primarily on catch history alone (CMSY approach) performed only 25% better than a random guess and assigned stocks to the wrong FAO status category (underfished, $B/B_{MSY} \geq 1.2$; maximally sustainably fished, $0.8 \leq B/B_{MSY} < 1.2$; overfished, $B/B_{MSY} < 0.8$) 57% of the time. Bouch et al. (2021) compared CMSY evaluations to ICES stock status determination from 17 assessed stocks from the Celtic, Irish, and North Seas and found that CMSY predicted higher relative F (11 of 17 stocks) and lower relative biomass (12 of 17 stocks). The simplicity of catch-only methods is appealing; however, applying them to obtain the status of individual stocks can provide highly imprecise and biased results (Ovando et al. 2021). Catch is a function of catchability, effort, and biomass, so using only data on catch to infer biomass trends requires assumptions of constant catchability and constant effort. Each unique combination of r and K, along with a model and catch history, produces a specific stock status at each time step. Since a prior is also provided for recent stock depletion, two priors are effectively defined in the Catch-MSY method (Ovando et al. 2021; 2022). This can result in a situation where an appropriate prior of recent depletion is provided but the priors for population biology (r and K) that appear to be independent of stock status strongly influence the posterior distribution of stock status (Ovando et al. 2021). Ovando et al. (2021) do however propose a method for tuning the r-K priors to have them be consistent with the explicitly defined depletion prior.

Several attempts of validating mechanistic catch-only methods have consistently found them to perform poorly (Free et al., 2020; Pons et al., 2020; Bouch et al., 2021; Sharma et al., 2021). For example, the evaluation of Bouch et al. (2021) of the method on data-rich stocks found that performance of the CMSY assessment highly depends on how well the ranges for the biomass priors match up with the "true" (data-rich) assessment. An analysis was conducted by Bouch et al. (2021) to assess the sensitivity of the priors to a 10% change. The authors found that the prior on upper depletion at the start and end of the time series had the greatest influence on B/B_{MSY}. Froese et al. (2022) point out that if the range of values of B/K used as a prior are too wide then it is telling the model that it is equally likely that the stock is collapsed vs. unexploited (which is unlikely).

Pons et al. (2020) found that CMSY performs poorly for stocks that are actively being managed where F has decreased in recent years and higher catches are interpreted as higher F instead of successful management. Sharma et al. (2021) found that CMSY methods performed well when reliable estimates of initial depletion were available but estimates of stock status were biased when using default priors provided in the CMSY software. Sharma et al. (2021) suggest that CMSY may be suitable for categorizing data-poor stocks in aggregate, given that the overall biases would be low across a number of stocks, but that catch-only methods are not a substitute for a data-rich assessment.

Based on the results from several attempts to validate CMSY methods, Ovando et al. (2021) suggest that improvements to the understanding of the stock status of global fisheries will come from improving data collection and not the development of data-limited methods alone. Ovando et al. (2022) showed that the catch time series does little to update the status predictions beyond the priors. The performance of the CMSY methods therefore depend on the reliability of the depletion priors (i.e., stock status in terms of relative depletion is known). If stock status in terms of depletion is known, an advantage of the CMSY methods could be that they can provide an estimate of stock status relative to B_{MSY}, however, DFO's PA Policy does not require the LRP to be defined in terms of B_{MSY}, and an LRP can be defined in terms of B_0 (K) without conversion to a ratio relative to B_{MSY}. As Bouch et al. (2020) states, "stock assessment is used to determine the depletion level of the stock, not rely on it being known beforehand". The relationship between B/B_0 to B/B_{MSY} is strongly based on r or steepness (h) of the SRR (Section 2.1.4) so an estimate of stock status relative to B_{MSY} based on depletion could be obtained without the use of CMSY methods.

C.3.2. Other Mechanistic Catch-only Methods

The optimized catch-only model (OCOM, Zhou et al. 2018) is a modified version of the CMSY approach for which the prior for r is derived from an estimate of the natural mortality rate and the prior for depletion is estimated using the Zhou et al. (2017) catch-only boosted regression tree model. The method performed well for 14 fish stocks assessed using stock synthesis (Zhou et al. 2018).

Two other mechanistic approaches include the catch-only model with sampling-importance-resampling (COM-SIR; Vasconcellos and Cochrane, 2005) and the state-space catch-only model (SSCOM; Thorson et al. 2013). COM-SIR is a coupled harvest dynamics model in which biomass dynamics follow a Schaefer model and harvest follows a logistic model. A sampling-importance-resampling algorithm is used in the model fitting. SSCOM is a coupled harvest dynamics model that estimates dynamics in biomass and harvest based on the catch time series, priors of r, the maximum rate of increase of fishing effort, and the magnitude of stochasticity for variation in effort, population dynamics, and fishing efficiency. The model is fit in a Bayesian state-space framework and was validated by simulation (Thorson et al. 2013).

OCOM was evaluated by Free et al. (2020) and COM-SIR and SSCOM were evaluated by Anderson et al. (2017) and all three methods were identified to provide imprecise and biased estimates of stock status when applied to data-rich assessments from the RAM Legacy Database and a set of simulated stocks.

Two Simple Stock Synthesis (SSS) methods have been developed for catch-only scenarios to mimic a depletion-based stock reduction analysis (DBSRA) (Cope 2013). SSS-MC uses a Monte Carlo approach like DBSRA by drawing values from probability distributions of input parameters (M, h, and depletion) and estimates R_0 and repeats this many times to obtain probability distributions for model outputs. SS-MCMC uses the DBSRA distributions as prior distributions for the parameters and a Markov chain Monte Carlo (MCMC) approach to calculate posterior distributions for all model outputs. SS-MCMC estimates M, h, R_0, while fitting to an artificial abundance survey representing stock depletion with an error distribution equivalent to the depletion prior used in the DB-SRA (Cope 2013). The SSS models require 4 input parameters (priors): distribution of relative depletion in the current year, distribution of M, distributions of F_{MSY}/M, and distributions of B_{MSY}/B_0 (these ratios are how productivity/h is incorporated).

Pons et al. (2020) evaluated the bias and precision of SSS and DBSRA and two versions and the CMSY method in a simulation study (simulation setup described in Section 2.11.2). SSS generally performed best among the catch-only methods, but it did underestimate harvest rates for medium-lived and short-lived species at high stock sizes. DBSRA was the most precise among the catch-only methods, but it generally underestimated harvest rates (Pons et al. 2020). The CMSY methods were the most biased and imprecise among the four catch-only methods evaluated.