# WORKSHOP ON ICES REFERENCE POINTS (WKREF1)

Please note: the list of authors was updated in January 2022.

VOLUME 4 | ISSUE 2

ICES SCIENTIFIC REPORTS

RAPPORTS SCIENTIFIQUES DU CIEM

ICES INTERNATIONAL COUNCIL FOR THE EXPLORATION OF THE SEA
CIEM CONSEIL INTERNATIONAL POUR L'EXPLORATION DE LA MER

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

H.C. Andersens Boulevard 44-46
DK-1553 Copenhagen V
Denmark
Telephone (+45) 33 38 67 00
Telefax (+45) 33 93 42 15
www.ices.dk
info@ices.dk

ISSN number: 2618-1371

# ICES Scientific Reports

Volume 4 | Issue 2

## WORKSHOP ON ICES REFERENCE POINTS (WKREF1)

## Recommended format for purpose of citation:

ICES. 2022. Workshop on ICES reference points (WKREF1).
ICES Scientific Reports. 4:2. 70 pp. http://doi.org/10.17895/ices.pub.9822

## Editors

Massimiliano Cardinale • Henning Winker

## Authors

Christoffer Moesgaard Albertsen • Valerio Bartolino • Mikaela Bergenius Nord • Santiago Cerviño Mikael van Deurs • Gjert Endre Dingsør •Daniel Duplisea • Johanna Fall • Dorleta Garcia • David Gilljam Nicolas Goñi • Michaël Gras • Tomas Gröhsler • Kristiina Hommik • Jan Horbowy • Daniel Howell • Leire Ibaibarriaga • Pekka Jounela • Laurence Kell • Christoph Konrad • Cecilie Kvamme • Debra Lambert • Jean-Baptiste Lecomte  Colm Lordan • Johan Lövgren • Patrick Lynch • Francesco Masnadi • Richard D. Methot • Tanja Miethe  David Miller • Cóilín Minto • Iago Mosqueira • Sofie Nimmegeers • José De Oliveira • Alessandro Orio  Martin Pastoors • David Reid • Rishi Sharma • Andreia Silva • John Simmonds • Henrik Sparholt • Sven Stoetera • Marc Taylor • Vanessa Trijoulet • Andrés Uriarte • Lies Vansteenbrugge • Sindre Vatnehol  Ching Villanueva • Laura Wise

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

# Contents

# i Executive summary


The ICES Workshop on ICES reference points (WKREF1) was tasked to provide a thorough review of the ICES reference points system as a basis to re-evaluate the process for estimating, updating and communicating reference points in the context of the ICES advice. As part of the preparation leading to WKREF1 a large database of the most recent assessment outputs for 78 Category 1 stocks were collated in the form `FLStock` objects, which formed the basis for several components of the presented analyses.

The first part of the meeting involved a detailed overview of the history and basis of the ICES references points system, which was aligned with the results of an empirical review of the procedures and choices made to derive ICES reference points for category 1 stocks. The ICES procedures were then contrasted with those used in the USA, Canada, New Zealand and across tuna Regional Fishery Management Organizations. A limitation in terms of transparency of the ICES procedures is a lack of complete documentation of the settings used for deriving reference points using, e.g., the EQSIM software. In comparison to other international standards, the main differences identified include the absence of a target biomass reference point and inconsistent estimates of the limit biomass reference point $B_{lim}$, which is estimated to be below 10% of the unfished biomass (B0) for a high proportion of analysed stocks (around 50%). In addition, an important difference is that direct estimates of $F_{MSY}$ are used in ICES (which can be unreliable), whereas elsewhere $F_{MSY}$ is often replaced by more conservative biological proxies, such as Fspr% and FB%.

The second part focused on the robustness evaluation of the current ICES reference point system. Work presented included examples that demonstrate differences in reference point estimates between standard ICES procedures (EQSIM) and full Management Strategy Evaluation (MSE) simulations, and a study that highlighted the considerable uncertainty in estimating $B_{lim}$, with guidance on how to quantify uncertainty depending on length and contrast in the time-series. A large simulation experiment was conducted by applying a short-cut MSE approach to 68 Category 1 stocks, which revealed that the ICES MSY advice rule is the least robust of all tested generic approaches if assumptions about the typically highly uncertain stock recruitment relationship are violated. This led to poor performance of the ICES MSY advice rule associated with the lowest long-term yields, highest risks to fall below limit reference points and lowest probabilities of attaining biomass levels at MSY. Contributing factors to the poor performance were combinations of comparably high $F_{MSY}$ estimates and low $B_{lim}$ (<10% $B_0$) and thus low MSY $B_{trigger}$ (<14% $B_0$) values. In cases where the precautionary FP.05 was invoked (mostly $B_{lim}$ > 10% $B_0$), performance improved notably. Results from a backtest using hindcasting with forecasts of 1-5 years reinforced the need to re-estimate reference points regularly at benchmark assessments, and that short-term forecasts should not exceed a three-year time span to account for time-varying biological traits. Finally, presented work highlighted the advantages (in terms of high consistency and accuracy) of estimating the stock-recruitment relationship or reference points internally in the assessment model, which was illustrated for Stock Synthesis and SAM, and supported by a comprehensive simulation study.

The key recommendations of WKREF1 were to: i) revise and simplify how $B_{lim}$ is derived. An absolute $B_{lim}$ should only be specified empirically in cases where there is sufficient contrast in the stock-recruit data to estimate a well-defined break-point. Alternatively, it is suggested that $B_{lim}$ should be determined as a plausible ratio of $B_0$ based on biological principles and the life history of the stock (e.g. 10-25% $B_0$ depending on the stocks characteristic; Section 7). ; ii) $F_{P.05}$ should be calculated without $B_{trigger}$; iii) to use biological proxies ($F_{brp}$) for deriving $F_{MSY}$, and the resultant $F_{MSY}$ proxy must not exceed $F_{P.05}$; iv) to report biomass target ($B_{trg}$) that corresponds to the $F_{MSY}$

proxy; and v) to set $B_{trigger}$ as either a fraction of $B_{trg}$ or multiplier of $B_{lim}$. Specifications of setting reference points (e.g. $F_{brp}$) or $B_{trigger}$ should be informed through further simulation testing to be presented at WKREF2.

# ii Expert group information

| | |
|---|---|
| **Expert group name** | Workshop on ICES reference points (WKREF1) |
| **Expert group cycle** | Annual |
| **Year cycle started** | 2021 |
| **Reporting year in cycle** | 1/1 |
| **Chairs** | Massimiliano Cardinale, Sweden |
| | Henning Winker, JRC-EC |
| **Meeting venue and dates** | 2-4 November 2021, Online meeting |

# Opening of the meeting

The ICES Workshop on ICES reference points (WKREF1) was held online, on 2–4 November 2021. The list of participants and contact details are given in Annex 2. The chairs, Massimiliano Cardinale (Sweden) and Henning Winker (Joint Research Centre – European Commission) welcomed the participants and highlighted the variety of Terms of References (ToRs). The draft agenda was presented and ToRs for the meeting (Section 1) were discussed. The Agenda was agreed and responsibility for individual tasks distributed among individuals.

# 1 Terms of Reference

The Workshop on ICES reference points (WKREF1) chaired by Massimiliano Cardinale, Sweden and Henning Winker, JRC-EC, will meet as a hybrid meeting online and in ICES, Denmark, November 2-4 2021 to:

(a) Review the current limit, trigger and target reference point estimation procedures for both biomass and fishing mortality used by ICES, and identify limitations and inconsistencies considering international best practice, the precautionary approach and international policies and legislation.

(b) Evaluate the robustness, consistency and plausibility of limit and target reference points in relation to current ICES Advice Rule in comparison to alternative approaches.

(c) Explore alternative methods that can better account for stock dynamics, biological realism and productivity drivers in reference point estimations under climate and environmental uncertainties.

(d) Consider appropriate methods of propagating model, estimation and process error uncertainties in the estimation of reference points.

(e) Propose candidate methods to address the emerging issues identified under (a) – (d).

WKREF will report by 29 November 2021 for the attention of the ACOM Committee.

**Supporting information**

| Priority | High |
|---|---|
| Scientific justification | ICES refers to two types of reference points when providing fisheries advice for category 1 stocks: precautionary approach (PA) reference points and maximum sustainable yield (MSY) reference points. The PA reference points are used when assessing the state of stocks and their exploitation relative to the precautionary approach objectives. The MSY reference points used in the advice rule (AR) applied by ICES are aimed at producing advice consistent with the objective of achieving MSY.<br><br>The PA reference points and the methods for estimating values were developed between 2001 and 2003 leading to the report of the Study Group on Precautionary Reference Points for Advice on Fishery Management in 2003 (ICES, 2003). Subsequently ICES was requested provide advice consistent with the objective of achieving<br>maximum sustainable yield (MSY). The guidance for estimation of MSY reference points has evolved based on the findings of four workshops: WKMSYREF2, (ICES, 2014), WKMSYREF3 (ICES, 2015), WKMSYREF4 (ICES, 2016), WKGMSE3 (ICES, 2020a) and Rindorf et al. (2017).<br><br>Recent ICES workshops (WKREBUILD, WKGMSE3 and WKRP-CHANGE) have discussed and made recommendations related to ICES reference points. The current procedures for estimating reference points |

are complex and difficult to communicate internally and externally. In recent years, a number of issues have been identified and these include:

- Should the AR be part of the of reference points estimation process? (because it assumes managers will always apply the AR)
- Conceptually $F_{P.05}$ without the advice rule is more consistent with the previous definition of $F_{pa}$ whereas ACOM decided to use $F_{P.05}$ with AR as $F_{pa}$.
- The rationale of capping $F_{MSY}$ and ranges with $F_{pa}$ or $F_{P.05}$ when also using the AR has been questioned.
- Proximity of $B_{lim}$ and MSY $B_{trigger}$ for some stocks impacting on AR.
- Which is the role of density dependence on reference point estimation.
- The evidence base for $B_{lim}$ in some stocks has been questioned.
- Inconsistency between EQSIM and MSEs.
- Inconsistency between how we assess risk in MSEs and reference point estimation.
- Stocks where $F_{lim} < F_{P.05}$ ($F_{pa}$)
- Subjectivity in the definition and use of S-R types for reference point estimation.
- Using segmented regression with a break point at $B_{loss}$ for Type 5 and 6 stocks.
- Unclear guidance for type 3 stocks and type 6 -high F stocks
- Whether all reference points are actually needed. Providing clear explanations what reference points are and how they are generated and what processes are included are important elements of a more consistent communication approach to reference points

A thorough review of ICES reference points is now needed and process for estimating, updating and communicating reference points needs to be re-evaluated in the context of the ICES advice rule and the recurrent advice that is currently been requested.

In relation to ToR b, the investigations should be carried out across stock assessment model platforms and include considerations of ratio-based reference points and implication of the stock recruitment relationship. Also, the implications of estimating reference points in the context of moving towards potential model ensemble based advice need to be considered.

| | |
|---|---|
| Resource requirements | One meeting room at ICES HQ with at least one breakout room and facilities for online participation. |
| Participants | Scientists with experience and interest in reference points definition and estimation procedures from inside and outside the ICES area. |
| Secretariat facilities | Secretariat administrative and scientific support. |
| Financial | No financial implications. |
| Linkages to advisory committees | The results of this work will directly feed the ICES advisory process |

| | |
|---|---|
| Linkages to other committees or groups | HAWG, WKGMSE3, WGWIDE, WGBFAS, WGCSE, WGNSSK, NWWG, AFWG, WGHANSA |
| Linkages to other organ zations | All advice recipients having an interest in ICES reference points. |

# 2 Scope of the workshop

The workshop included the requirement to review the current limit, trigger and target reference point estimation procedures for both biomass and fishing mortality used by ICES. Also, it was requested to evaluate the robustness, consistency and plausibility of limit and target reference points in relation to current ICES Advice Rule and in comparison to alternative approaches that can better account for stock dynamics, biological realism and productivity drivers in reference point estimations. Reference point estimation should have been evaluated under climate and environmental uncertainties and time-varying processes. Finally, the workshop should propose candidate methods to address the emerging issues identified under ToR (a) – (d).

## 2.1 Glossary

$B_{lim}$: A deterministic biomass limit below which a stock is considered to have reduced reproductive capacity. For stocks where quantitative information is available, a reference point $B_{lim}$ may be identified as the stock size below which there is a high risk of reduced recruitment.

$B_{loss}$: It is the lowest observed SSB in the assessment time series and commonly used as a proxy for $B_{lim}$ (i.e. Type 5 within the current ICES advice framework).

$B_{pa}$: A precautionary safety margin incorporating the uncertainty in ICES stock estimates leads to a precautionary reference point $B_{pa}$, which is a biomass reference point designed to have a low probability of being below $B_{lim}$. When the spawning-stock size is estimated to be above $B_{pa}$, the probability of impaired recruitment is expected to be low. $B_{pa}$ is estimated as a function of $B_{lim}$.

$F_{lim}$: The fishing mortality which in the long term will result in an average stock size at $B_{lim}$. Fishing at levels above $F_{lim}$ will result in a decline in the stock to levels below $B_{lim}$.

$F_{P.05}$: The fishing mortality that results in no more than 5% probability of bringing the spawning stock to below $B_{lim}$ in the long term.

$F_{pa}$: Same as $F_{P.05}$

$F_{MMY}$: The maximum medium yield $F_{MMY}$ denotes the fishing mortality that corresponds to the peak of the median landings yield curve derived from stochastic forward projections as is typically derived from the EQSIM software (i.e. "FMSYmedianL"). Within the ICES advice framework, the quantity $F_{MMY}$ is typically referred to as $F_{MSY}$. However, for FMMY to directly translate into FMSY as reported on the advice sheet, $F_{MMY}$ first requires meeting the condition that $F_{MMY} < F_{P.05}$ in accordance with precautionary principle. For the purpose of this report a clearer definition was therefore needed to separate the initial estimate of $F_{MSY}$, here $F_{MMY}$, from the final advice for $F_{MSY}$.

$F_{MSY}$: Within the ICES advice framework $F_{MSY}$ is specified as $F_{MSY} = min(F_{P.05}, F_{MMY})$. Within an international or operating model (simulation) context, $F_{MSY}$ is referred to as a biological reference point that specifies the fishing mortality rate that, if applied constantly, would result in an average catch corresponding to the Maximum Sustainable Yield (MSY) and an average biomass corresponding to $B_{MSY}$.

MSY $B_{trigger}$: MSY $B_{trigger}$ is the parameter in ICES MSY framework which triggers advice on reducing fishing mortality relative to $F_{MSY}$. MSY $B_{trigger}$ is considered the lower bound of SSB fluctuation (fifth percentile of the $B_{MSY}$ estimate) when fished at $F_{MSY}$, but is set for a large majority of stocks equal to $B_{pa}$.

$B_{trigger}$: Generalization of the MSY $B_{trigger}$, which can differ in the way it is specified.

$B_{MSY}$: It is the expected average biomass if the stock is exploited at $F_{MSY}$, but currently not reported in ICES.

$B_0$: In age-structured models, B0 is the unfished spawning biomass that is given by the product of virgin recruitment $R_0$ (implicit to the stock recruitment relationship) and the unfished spawning biomass-per-recruit ($SPR_0$) being a function of weight-at-age, maturity-at-age and natural mortality. Like $B_{MSY}$, it is therefore an implicit property of any age-structured model for which a SRR is estimated or assumed, but currently not reported in ICES.

$SB_{\%}$: The percentage spawning stock biomass of the unfished $B_0$ (e.g. $B_{40}$)

MSY Proxies: Analytical proxies for $B_{MSY}$, $F_{MSY}$ and MSY are quantitative surrogates that can be used if direct estimation is not possible or the estimates are not considered reliable.

$F_{max}$: The fishing mortality at which the yield-per-recruit is maximized. $F_{max}$ remains relevant to the ICES advice rule in many cases where segmented regression is assumed for stock recruitment relationship, because $F_{MMY}$ as the direct estimate of $F_{MSY}$ is the same as $F_{max}$ on a yield-per-recruit curve for the range of SBB > $B_{lim}$.

$F_{0.1}$: The fishing mortality at which the slope of the yield-per-recruit curve is 10% of that at the origin

$F_{spr\%}$: The fishing mortality at which the spawner-biomass-per-recruit (SPR) is, e.g. 40%, of its unexploited level $SPR_0$ (a common range is $F_{spr30}$ -$F_{spr50}$).

$F_{B\%}$: The fishing mortality at which the spawning stock biomass (SSB) is e.g. 40% of its unexploited level at $B_0$, i.e. $F_{B40}$

$F_{brp}$: Biological reference point proxy of $F_{MSY}$ (e.g. $F_{spr\%}$ and $F_{B\%}$)

## 2.2 Database

The database collated included 78 stocks which are assessed as category 1 by ICES in 2020 and 2021. Of those, 7 stocks were excluded as MSY reference points are undefined (i.e. cod.27.1-2coastN, cod.27.24-32, san.sa.1r, san.sa.2r, san.sa.3r, san.sa.4, spr.27.3a4 and reb.27.1-2). her.27.6a7bc was also excluded as, although it is assessed by SAM, it has been treated as category 3 since 2019.

For the following stocks it was not possible to conform the original assessment data into the FLR (Kell *et al*., 2007) framework (1), a complete dataset was not delivered or/and was not available (2) or no data was supplied before the deadline (3). These stocks are ghl.27.1-2 (2), dgs.27 (3), bli.27.5b67 (2) and nop.27.3a4 (1). For nop.27.3a4, the stock was anyhow excluded from the analysis as MSY reference points are undefined. Finally, two capelin stocks (cap.27.1-2 and cap.27.2a514), which are assessed as category 1 by ICES in 2021 were not requested as MSY reference points are undefined.

For each stock, an FLR object of class FLStock was created, which contained all input data and biology used in the assessment (except survey indices) and output data from the model. For all stocks, zeros values, when present, were replaced by a small constant to avoid issues with the simulations. When discards or landing weight of a particular age class were missing, they were replaced by the corresponding weight at age of the catches. For mac.27.nea, weight at the age of age-0 individuals was missing and it was replaced by 0.063 kg as described in the WGWIDE report. For stocks for which target fishing mortality is expressed as harvest rate (i.e. cod.27.5a, had.27.5a and pok.27.5a), target harvest rate was translated into $F_{bar}$.

Additional information was appended to each of the FLR stock objects, namely life history parameters obtained from FishLife (Thorson, 2020; [https://github.com/James-Thorson-NOAA/FishLife](https://github.com/James-Thorson-NOAA/FishLife)), the year of the assessment and the model used. Finally, also order, family, species resilience, species grouping, geographical area *sensu* ICES and the expert Working Group that conducted the assessment were collated in a separate database (i.e. not directly in the FLR stock objects).

Reference points were updated for all stocks using the ICES official most recent estimated values. Reference points for pil.27.8c9a used in the analysis refers to medium productivity (ICES 2021c).

**Table 2.1. Summary by region of all stocks included in the database used in the analyses.**

| Species | Baltic Sea | Barents Sea | Bay of Biscay and Iberian coast | Celtic Seas | Faroes | Greater North Sea | Green-land Sea | Ice-landic waters | Oceanic North East At-lantic |
|---|---|---|---|---|---|---|---|---|---|
| *Ammodytes marinus* | | | | | | 4 | | | |
| *Argentina silus* | | | | | | | | 2 | |
| *Brosme brosme* | | | | | | | | 1 | |
| *Clupea harengus* | 4 | 1 | | 3 | | 1 | | 1 | |
| *Dicentrarchus labrax* | | | | 1 | | 1 | | | |
| *Gadus morhua* | 2 | 2 | | 2 | 1 | 1 | 2 | 1 | |
| *Glyptocephalus cyno-glossus* | | | | | | 1 | | | |
| *Lepidorhombus boscii* | | | 1 | | | | | | |
| *Lepidorhombus whiffi-agonis* | | | 2 | | | | | | |
| *Lophius piscatorius* | | | 2 | | | | | | |
| *Melanogrammus ae-glefinus* | | 1 | | 3 | 1 | 1 | | 1 | |
| *Merlangius merlangus* | | | | 3 | | 1 | | | |
| *Merluccius merluccius* | | | | | | 1 | | | |
| *Micromesistius poutassou* | | | | | | | | | 1 |
| *Molva molva* | | | | | 1 | | | 1 | |
| *Pandalus borealis* | | | | | | 1 | | | |
| *Pleuronectes platessa* | 1 | | | 1 | | 2 | | | |
| *Pollachius virens* | | 1 | | | 1 | 1 | | 1 | |
| *Sebastes norvegicus* | | 1 | | | | | | 1 | |

| Species | Baltic Sea | Barents Sea | Bay of Biscay and Iberian coast | Celtic Seas | Faroes | Greater North Sea | Green-land Sea | Ice-landic waters | Oceanic North East At-lantic |
|---|---|---|---|---|---|---|---|---|---|
| *Sebastes mentella* | | 1 | | | | | | | |
| *Sardina pilchardus* | | | 2 | | | | | | |
| *Scomber scombrus* | | | | | | | | | 1 |
| *Scophthalmus maximus* | | | | | | 1 | | | |
| *Solea solea* | 1 | | 1 | 2 | | 3 | | | |
| *Sprattus sprattus* | 1 | | | | | 1 | | | |
| *Trachurus trachurus* | | | 1 | | | 1 | | | |
| | 9 | 7 | 9 | 15 | 4 | 21 | 2 | 9 | 2 |

**Table 2.2. Summary by region and assessment model used of all stocks included in the database used in the analyses.**

| Area | a4a | AAP | AMISH | ASAP | Bayes-ian | Gadget | NFT-ADAPT | SAM | SCAA | SMS | SS3 | TSA | XSA |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Baltic Sea | | | | | | | | 4 | | | 2 | | 3 |
| Barents Sea | | | | | | 1 | | 5 | 1 | | | | |
| Bay of Biscay and Iberian coast | 1 | | 1 | | 1 | | | | | | 3 | | 3 |
| Celtic Seas | | | | 3 | | | | 9 | | | 1 | | 2 |
| Faroes | | | | | | | | 4 | | | | | |
| Greater North Sea | | 3 | | | | | | 7 | | 5 | 4 | 1 | 1 |
| Greenland Sea | | | | | | | | 2 | | | | | |
| Icelandic waters | | | | | | 4 | 1 | 1 | 3 | | | | |
| Oceanic North East Atlantic | | | | | | | | 2 | | | | | |
| | 1 | 3 | 1 | 3 | 1 | 5 | 1 | 34 | 4 | 5 | 10 | 1 | 9 |

## 2.3 Estimation of $B_0$ and $B_{MSY}$ in EQSIM

For each stock, we run a standardized EQSIM (ICES 2017; https://github.com/wgmg/msy) simulation with the intent to replicate the most common configuration within ICES for the estimation of MSY reference point. EQSIM (stochastic equilibrium reference point software) provides MSY reference points based on the equilibrium distribution of stochastic projections. Productivity parameters (i.e. year vectors for natural mortality, weights-at-age, maturity-at-age, and selectivity) are resampled at random from the last years of the assessment. Recruitments are resampled from their predictive distribution which is based on parametric models fitted to the full time-series provided. Uncertainty in the stock–recruitment relationship (SRR) can be taken into account by applying model averaging using smooth AIC weights (Buckland *et al.*, 1997) although in many stocks the SRR is taken to be just a single SRR. The segmented regression is the most commonly used SRR. For example, WKMSYREF4 (ICES, 2015) used mostly only segmented regression S-R functions in order to be compatible with the precautionary considerations. Thus, for WKREF1 we used only segmented regression S-R functions in the simulations with the breakpoint set at ICES $B_{lim}$. based on the 2021 advice and $F_{MSY}$ was set equal to the ICES official values. EQSIM was run without and with harvest control rule and for the latter the current MSY $B_{trigger}$ was used as input. Productivity parameters and selectivity were set as the average of the last 3 years for all stocks.

For the assessment error in the advisory year and the autocorrelation in assessment error in the advisory year, ICES default values (i.e. 0.212 and 0.423) were used (ICES 2015). Simulations were run for 200 years with the last 50 years retained to compute equilibrium values from. Autocorrelation of recruitment was used in all EQSIM simulations. The key quantities $B_0$ and $B_{MSY}$ were derived from the EQSIM simulations corresponding to the stock spawning biomass (*SSB*)at F=0 and F = $F_{MSY}$, respectively.

In general, a Beverton-Holt SRR produces almost always higher estimates of $R_0$ thus $B_0$ than a segmented regression. Only in very few cases where a weighted average of the Ricker and the segmented regression is used, $B_0$ may be lower than estimated from hockey-stick alone (c.f. ICES 2014; WKMSYREF3), whereas weighted averaging of all three models lead to higher $B_0$ values as can be inferred from the analyses presented in WKREF2 (ICES, 2014). Although not exact, the estimates of $B_0$ derived here can generally be treated as representative for the majority of stocks for which current advice is based on the segmented regression, and else, should be interpreted as a conservative (i.e. lower than nearly any other functional SRR used in ICES) estimate of $B_0$ than a computed value from the actual benchmark, with very few exceptions.

# 3 Review the current limit, trigger and target reference point estimation procedures for both biomass and fishing mortality used by ICES, and identify limitations and inconsistencies considering international best practice, the precautionary approach and international policies and legislation.

## 3.1 On the need to revise reference points

*by Colm Lordan and David Miller*

The ecosystem approach is the central tenet that governs how ICES provides independent advice on the management of human activities in our seas and oceans. Through its advice, ICES strives to advance and share scientific understanding of marine ecosystems and the services they provide to meet conservation, management, and sustainability goals. ICES advice is diverse and requested by multiple different requestors. The general framework and principles governing all ICES advice is set out here in ICES (2020a).

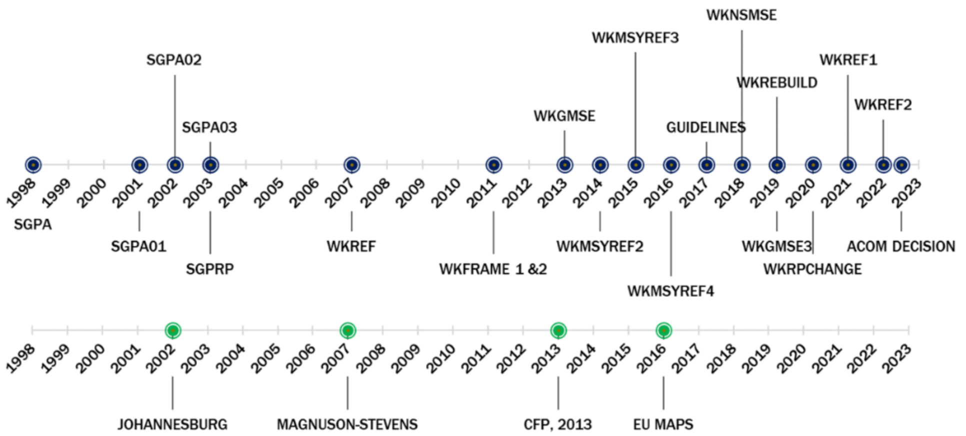


**Figure 3.1. The timeline that gives key developments in the evolution of ICES PA and MSY reference points.**

The thinking and process around ICES precautionary and MSY reference has developed and evolved over time strongly influenced by policy needs and drivers but also the availability of tools to estimate reference points in a consistent way. SGPA in 1998 defined $B_{lim}$ as the biomass “below which recruitment becomes impaired or the dynamics of the stock are unknown, and that fishing mortality at $F_{lim}$ that will drive the spawning stock to that biomass limit” (ICES, 1998). The word ‘impaired’ is synonymous with the concept that on average recruitment becomes systematically reduced as biomass declines below a certain point. During the early 2000s the various SGPA meetings developed the thinking around precautionary reference points considerably (ICES, 2001, 2002, 2003a) culminating in SGPRP in 2003 which was the first systematic attempt to estimate PA reference points for most data rich ICES stocks (ICES, 2003b).

At that stage, ICES had already started to define stock-recruit (SR) Types, use Segmented Regressions to estimate break points. The definition of $B_{lim}$ was as “the SSB below which there is a substantial increase in the probability of obtaining reduced (or 'impaired') recruitment i.e. the estimate of $B_{lim}$ should be risk averse, so that when the stock is at $B_{lim}$ the probability that recruitment is substantially impaired is still small, but below $B_{lim}$ that probability increases."

In 2002, the Johannesburg Declaration of the World Summit on Sustainable Development (WSSD; UN, 2002), called for an ecosystem approach and rebuilding fisheries to maximum sustainable yield. In 2007, WKREF was established as a workshop on Limit and Target Reference Points (ICES, 2007). Various problems with limits and targets were identified and there was no consensus on a way forward. It was thought "that distance between $B_{pa}$ and $B_{lim}$ could take into account the uncertainty due to different regimes". WKREF concluded from the review of the scientific and management literature that Maximum Sustainable Yield (MSY) is a difficult concept for management purposes because it is difficult to assess, unstable over time and only applicable in a single species context. Single species MSY and $B_{MSY}$ will not work for predators and prey at the same time.

WKFRAME in 2010 and 2011 was tasked with drafting technical guidelines to assist ICES expert groups in the implementation of the ICES MSY framework for advice (ICES, 2011). WKFRAME discussed the role of $B_{trigger}$ and indicated "it should be selected as a biomass that is encountered with low probability if $F_{MSY}$ is implemented" and that "under MSY exploitation should be a property of the expected distribution of SSB". However, ensuring compatibility with the PA was also raised as an issue including the need to avoid $B_{lim}$ in the long term taking account of model error. At this stage generic tools such as PlotMSY, which used the assessment summary and sensitivity data (so called sum and sen files), became available to the ICES community. This was a major step forward to stochastically estimate YPR based reference points. The main limitation of PlotMSY was that different SR forms were not modelled.

Various ICES advice recipients developed strong policies to implement an ecosystem and MSY approach in their fisheries management system. Within the EU, legal obligations to implement MSY management and establish multiannual plans reflecting the specificities of different fisheries based on the best available science were set out in the reformed CFP (EU, 2013). There were significant technical developments around management strategy evaluations (WKGMSE, ICES 2013, Punt et. al, 2015) and work on developing an ICES tool to estimate MSY reference points (EQSIM) began. After limited progress at WKMSYREF1, there were significant developments as EQSIM was more widely tested at WKMSYREF2 (ICES, 2014). A joint ICES/MYFISH (https://www.myfishproject.eu/) workshop WKMSYREF3 (ICES, 2015) systematically estimated MSY reference points and $F_{MSY}$ ranges for the North Sea and Baltic stocks to address a special request from the EU for MSY ranges (ICES, 2014b, ICES, 2015). WKMSYREF4 developed the approach further and estimated MSY ranges for demersal stocks in western waters (ICES, 2016a, ICES, 2016b). The emerging 5-step procedure for estimating reference points was strongly linked to the advice framework and the need to input the ICES MSY advice rule (AR) to ensure it is consistent with the ICES Precautionary Approach. The ICES technical guidelines to estimated "fisheries management reference points for category 1 and 2" was published in 2017 (ICES, 2017).

The presentation included a number of slides to explain the existing categorisation of stocks (Cat 1-2 MSY advice, Cat 3-6 PA advice), and the nature of hierarchy of the single stock advice (Management Plan, ICES MSY Approach, ICES PA). Currently, there are 25 stocks with management plan advice, 60 with ICES MSY approach advice, 17 with $F_{MSY}$ range advice (EU MAP), 4 with MSY approach including precautionary considerations, 138 stocks with precautionary advice, 7 stocks with status only advice and 7 stocks where no advice is provided.

A key consideration in defining target and limit reference points is how they are currently used and interpreted in advice products. The Fisheries and Ecosystem Overviews include an analysis of how many stocks are above and below MSY and PA reference points currently and also include plots of trends over time. Stock status in relation to reference points is also used for the

Marine Strategy Framework Directive (MFSD), Good Environmental Status (GES) D3 descriptor and globally for UN sustainability goals (Hilborn, 2020).

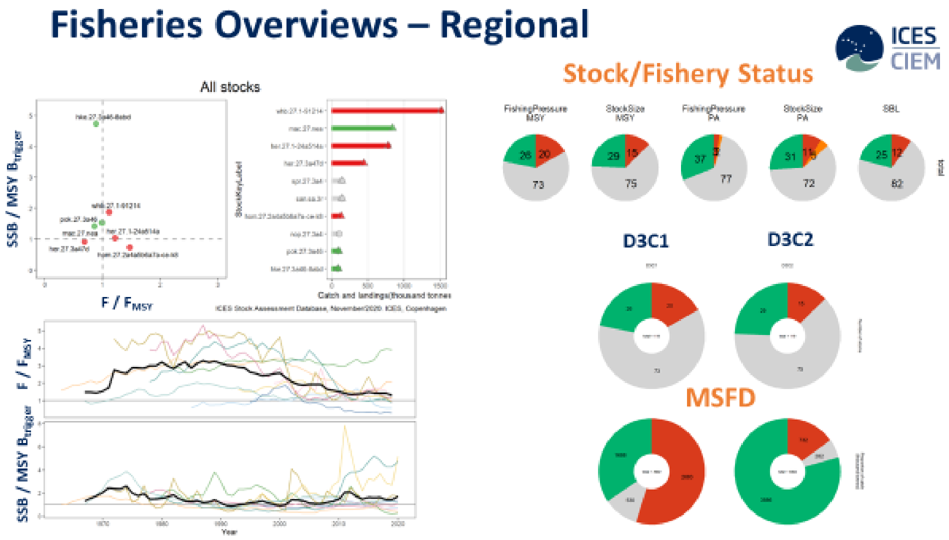


**Figure 3.2. Examples of how stock status in relation MSY and PA reference points are summarised in Fisheries and Ecosystem overviews at an ICES Ecoregion scale.**

A number of background documents were made available to WKREF1 outlining the international and national fisheries management policies on the SharePoint site (https://community.ices.dk/ExpertGroups/wkref1/2021_Meeting_Documents/02.%20Background%20documents/ICES%20International%20and%20national%20management%20objectives%20WITH%20LINKS.docx). Most advice recipients use similar terminology around the need to establish limit reference points such that "*Limit reference points set boundaries to constrain harvesting within safe biological limits so stocks can produce maximum sustainable yield*" and target reference points where "*Fishery management strategies shall ensure that target reference points are not exceeded on average*" (based originally on the UN Fish Stocks Agreement 1995 UNFSA). ICES gives advice based on requests from a range of requestors including governments, governmental agencies, RFMOs, commissions etc. The various Memoranda of Understanding (MoUs) and grant agreements were also made available to WKREF1 as background on the nature of the advice being requested (https://community.ices.dk/ExpertGroups/wkref1/2021_Meeting_Documents/02.%20Background%20documents/ICES%20MOUs-Agreements%20relevant%20to%20WKREF.docx ).

The presentation summarised the current ICES guidelines for estimating reference points (ICES, 2021a). This involves:

1. Identifying appropriate data (truncate time series or not).
2. Identifying stock type (6 different types described with different recommended actions).
3. Estimating biomass limit reference points.
4. Deriving PA reference points from limit reference points.
5. Estimate MSY reference points without and later with the AR (MSY Btrigger) to test for precautionarity.

The current guidelines (ICES, 2021a) are complex, convoluted and not always well understood or followed by assessment practitioners. While the documentation in the guidelines is reasonably good there is no documentation with EQSIM to help those at benchmarks with implementation and interpretation.

Since 2017, a number of ICES expert groups (WKNSMSE, WKREBUILD, WKGMSE3, WKMSEMAC, WKREPCHANGE), have identified challenges and suggested developments that need to be addressed by WKREF1 and WKREF2 (ICES, 2019, 2020bcd, 2021b). These were summarised and synthesised in the presentation. WKREF1 needs to consider the differences between target and limit reference points. How targets and limits might be used in the ICES MSY advice framework (which includes the advice rule). More generically, how reference points are used in management plan evaluations to define if various proposed harvest control rule are considered "precautionary and in accordance with the MSY approach". The final point was a call to action outlining what is now needed; simplification of methodology and terminology, better consistency, a generic approach applicable across multiple stocks and methods, while addressing PA/MSY needs, taking better account of changing productive drivers (growth, reproduction, recruitment, density dependence DD, survival). The main outputs should be clear recommendations to ACOM which is evidence based and a road map to implementation so that WKREF2 can develop user friendly guidelines and tools for the future.

## 3.2 An empirical review of current ICES reference point estimation procedures

*Paula Silvar-Viladomiu, Luke Batts, Cóilín Minto, and Colm Lordan*

We provide an extensive empirical review of procedures and settings used to derive current ICES category 1 reference points. A comprehensive database has been assembled for 79 stocks from all available literature. Reference point settings are visualised. Comparing across stocks illustrates both consistency and some inconsistencies in stock-recruit type and reliant technical basis. We recommend that reference point estimation procedures be consistently documented and made reproducible in a transparent framework.

ICES advice is governed by several international agreements (from UN and FAO), policies of ICES member countries (CFP of the EU, NMRA of Norway, RFLF of Russia, IFM of Iceland), and multinational and intergovernmental organizations (MSFD, AMMR), see general context of ICES advice. Currently, ICES approach to advice on fishing opportunities integrates the precautionary approach with the objective of achieving maximum sustainable yield.

Here, we provide an empirical review of the current framework with respect to WKREF1 ToR a) "Review the current limit, trigger and target reference point estimation procedures for both biomass and fishing mortality used by ICES, and identify limitations and inconsistencies considering best practice, the precautionary approach and international policies and legislation."

In the context of WKREF 1 ToR a), our goals are to:

assemble a database of reference point estimation settings for 79 category 1 stocks. All available documentation was reviewed including: benchmark and inter-benchmark reports, working group reports, special requests, expert group reports, and specific working documents. Data collated include: estimation metadata, estimation timelines, estimation framework, and settings, stock-recruit settings, assessment error settings, period settings, references, reference point framework, stock-recruit type, and technical basis, EQSIM settings, hitting precautionary bounds. Additional stock-recruit data were extracted via XML parsing.

R code was developed to clean the information as collated from the documents. This typically comprised grouping categories. Our methods comprise a visualisation of the reference point estimation data; and an additional calculation of the coefficient of variation of recruitment residuals plotted against the coefficient of variation of spawning stock biomass to determine consistency with stock-recruit typology guidelines.

## Simulation/estimation framework and timeline

EQSIM was the dominant estimation framework used for the majority of stocks where reference points were estimated (Figure 3.3). Management strategy evaluations were used for 11 stocks. Most reference points have been updated in the last 5 years with two stocks with long established reference points (Northeast Arctic capelin and cod). There are 4 stocks with recent estimates of $F_{MSY}$ but older estimates of $B_{lim}$ reflecting changes from $F_{MSY}$ to $F_{P.05}$.

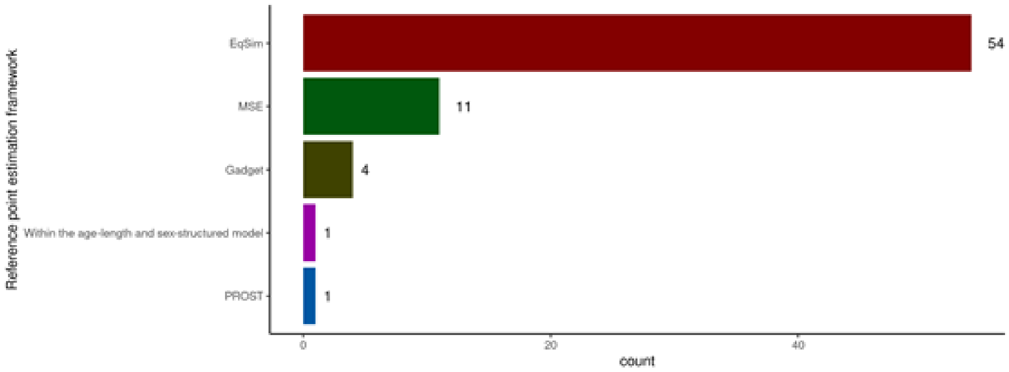


**Figure 3.3. Framework used to estimate reference points in ICES.**

## $B_{lim}$ technical basis

Recommended $B_{lim}$ technical basis depends on the SR typology according to the following guidelines. Tabulating reported stock-recruit typologies against the $B_{lim}$ technical basis demonstrates that for many stocks the SR type is not specified in the documentation. We subsequently categorised SR types on the basis of examination of the SR data and $B_{lim}$ to classify all but 4 stocks (Figure 3.4).

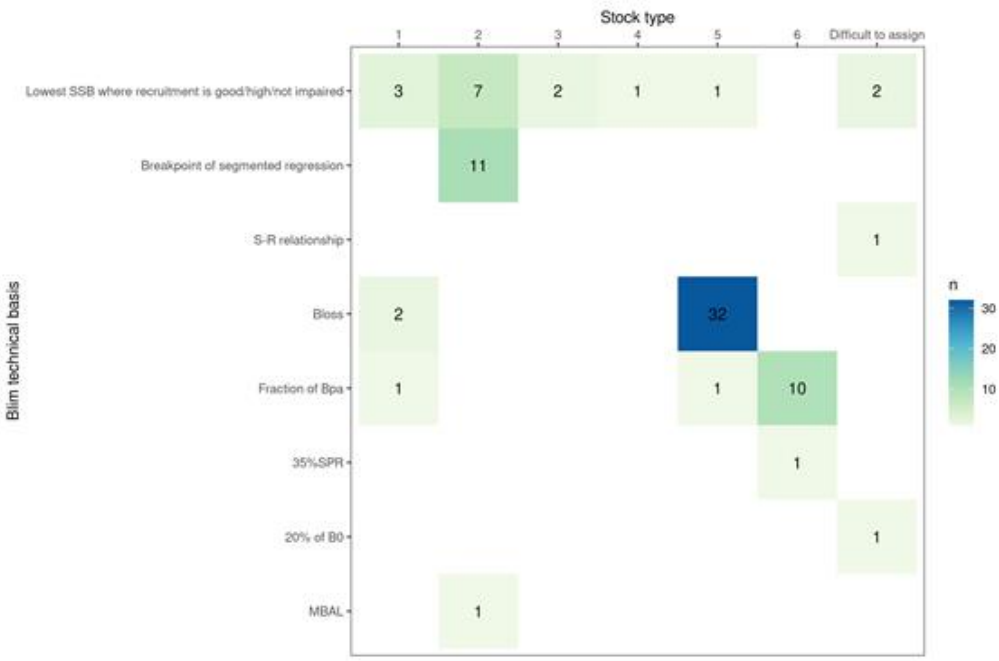


**Figure 3.4. Cross tabulation of reported stock-recruit typology and $B_{lim}$ technical basis.**

Plotting the observed variation in recruitment residuals versus variation in SSB (Figure 3.5) shows that some assigned typologies adhere well to their definitions (type 6 - narrow range of SSB), whereas there are examples of similar degrees of variation in SSB being categorised as wide or narrow ranges.

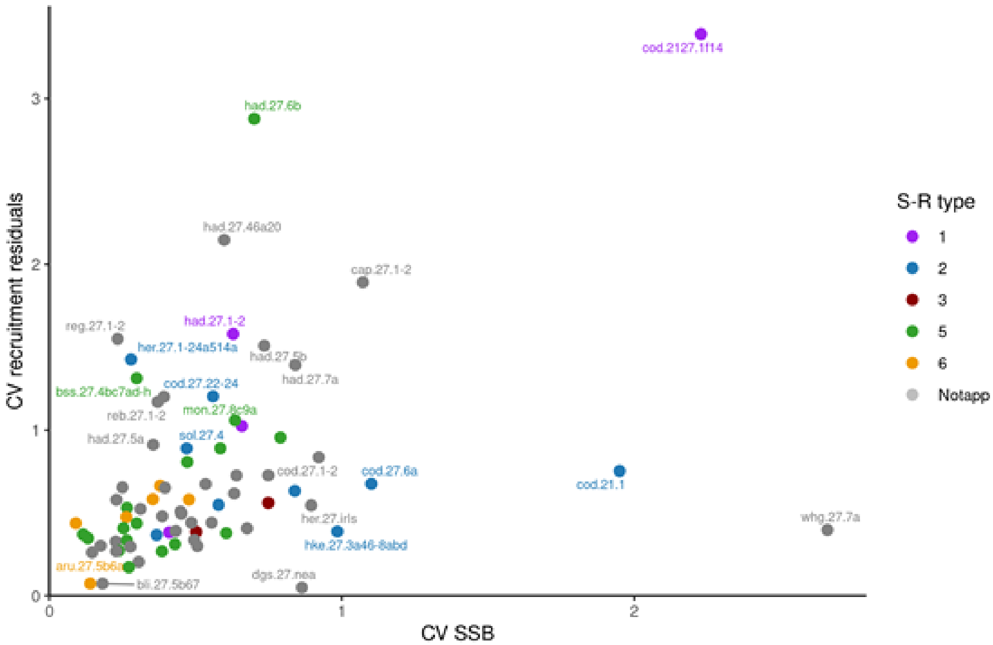


**Figure 3.5. Variation by SR type. y-axis shows the coefficient of variation of recruitment residuals (around a Beverton-Holt fit), x-axis shows coefficient of variation of SSB.**

### 3.2.1 Estimation settings

All stocks included process error in recruitment and most autocorrelation in recruitment. Default values were commonly used for assessment errors. Time windows for biological or selectivity values were typically 5 or 10 years. We note the default is the last 5 years in contrast to the guidelines that recommend 10 years.

### 3.2.2 Precautionary bounds, $F_{MSY}$ and reference point changes and impacts

Forty-three stocks did not hit the precautionary bounds while 12 stocks did. We have found that reference points can change relatively frequently with substantial changes between years. The main impact of changes is on $B_{lim}$ which changes MSY $B_{trigger}$ (Figure 3.6) as it is often defined as $B_{pa}$, which is often a multiple of $B_{lim}$. Changes in $B_{lim}$ do also impact $F_{MSY}$ where $F_{MSY}$ is set at $F_{P.05}$ due to a higher than 5% probability of hitting $B_{lim}$ (Figure 3.7).

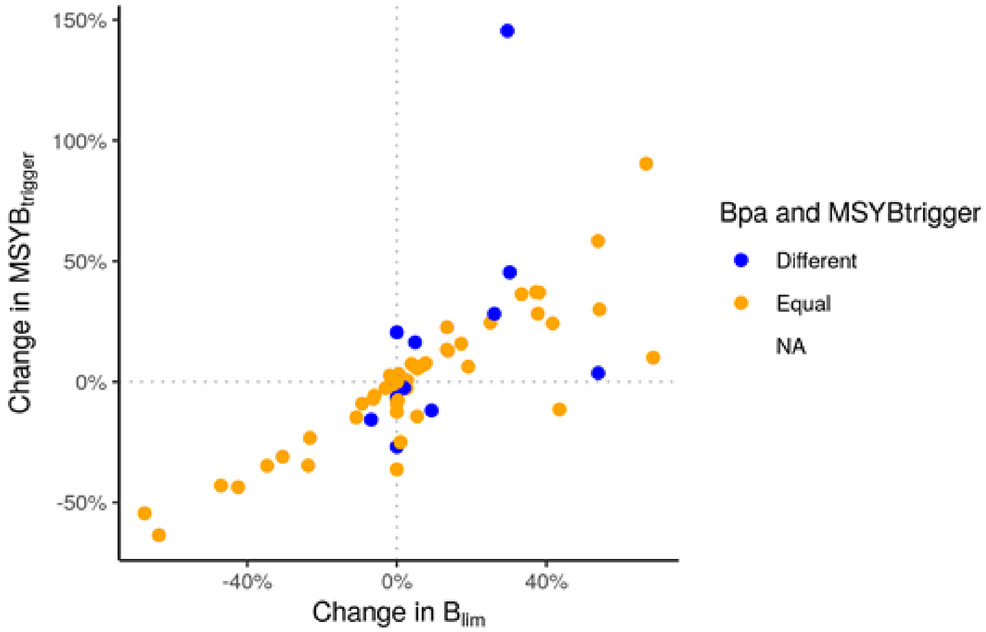


**Figure 3.6. Impact of changes in $B_{lim}$ on the ICES biomass trigger point MSY $B_{trigger}$.**

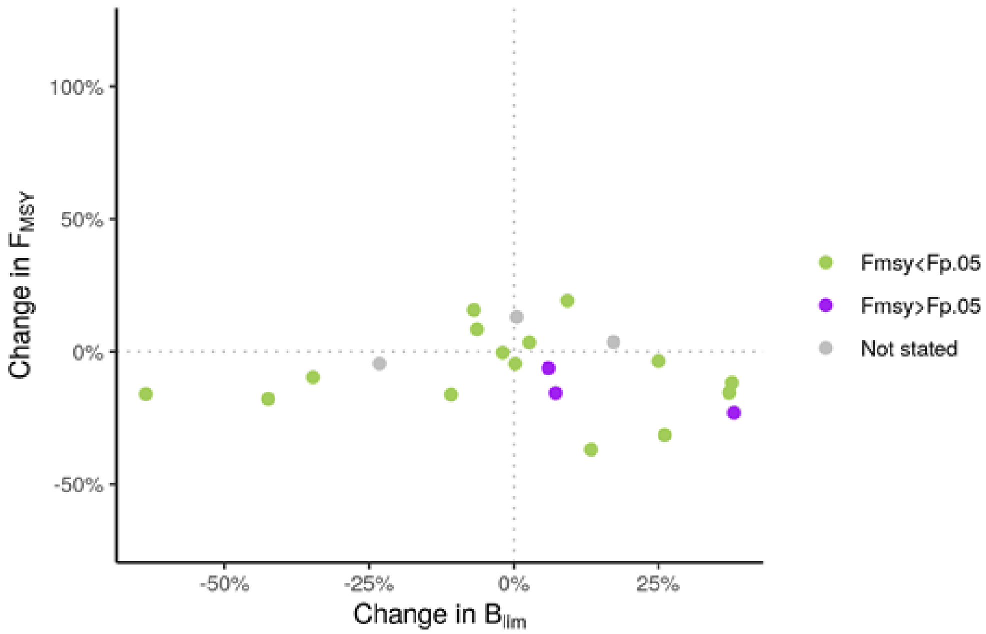


**Figure 3.7. Impact of changes in $B_{lim}$ on $F_{MSY}$.**

## 3.2.3 Discussion and conclusions

By extensively reviewing current ICES reference point estimation, we conclude that the current system incorporates many features of the precautionary approach as it pertains to recruitment overfishing. There is a need, however, for consistent and clear documentation of both general (cross-framework) and specific (EQSIM) decisions/settings made (as TAF is to assessments). As a community we should aim for retrospective reproducibility of full procedure (ability to replicate reference point estimation at any historical time point). There is important ambiguity in deciding SR types that becomes apparent when reviewing all stocks. The current framework is a classification one rather than a hypothesis-driven framework. As a minor note, current default settings should be aligned with guidelines regarding the choice of biological and selectivity years.

## 3.3 Reference Points in tunas RFMO's and General Principles on reference points; Putting ICES in context with these targets and limits

*by Rishi Sharma*

An overview of the general principles of precautionary management were presented. The key target and limit reference points (Figure 3.8) were presented based on MSY targets or Virgin Biomass size ($B_0$).

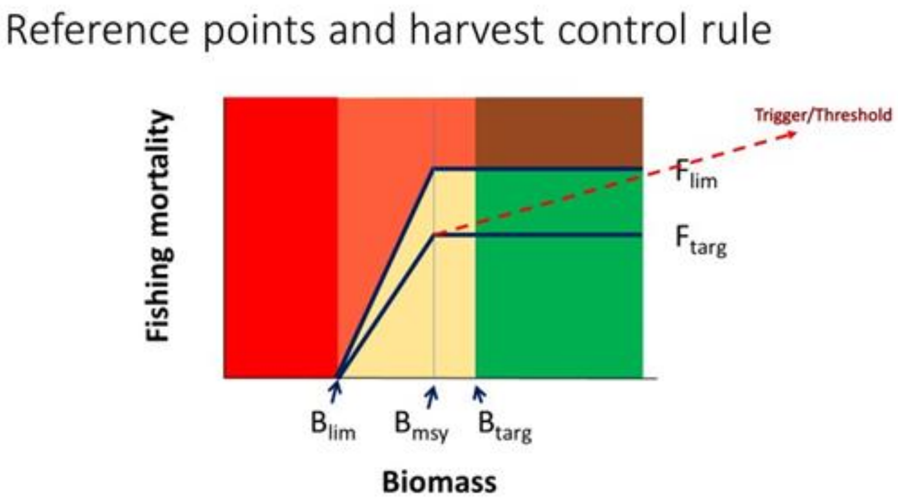


**Figure 3.8. Biomass reference points and limit and target fishing mortalities related to them, $B_{trigger}$ is the point at which management actions are taken in the fishery.**

Approaches pursued across 5 tunas RFMO's were discussed and all stemmed from a target based on $B_{MSY}$ targets, and limits were in most cases set of 0.4 $B_{MSY}$ for IOTC and ICCAT (assuming that we use a Schaeffer production function this corresponds to 0.2 $B_0$). CCSBT had a target set at $0.3B_0$, and a limit to $0.2B_0$. IATTC was the most risk prone and used an analytical solution based on where recruitment is reduced by 50%; analytically this is approximately 0.067 of $B_0$. All 4 RFMO's used these target and limits as achievable in a probabilistic sense (i.e. probability of exceeding target (x%), and probability exceeding limit (y%) based on multi-model ensembles to assess uncertainty. WCPFC used MSY based management targets as well, but had a more complicated design for limits based on how well steepness is known for the stock (Table 3.1). WCPFC agreed on a limit reference point of 20% $B_0$ at WCPFC9 in 2012 for all 'key' tuna species.

**Table 3.1. The table indicates how limits were set up for species managed under WCPFC.**

| Level | Condition | LRPs |
|---|---|---|
| Level 1 | A reliable estimate of steepness is available | FMSY and BMSY |
| Level 2 | Steepness is not known well, if at all, but the key biological (natural mortality, maturity) and fishery (selectivity) variables are reasonably well estimated. | FX%SPRo and either 20%SBo or 20%SBcurrent,F=0 |
| Level 3 | The key biological and fishery variables are not well estimated or understood. | 20%SBo or 20%SBcurrent,F=0 |

All 5 tRFMOs have these targets and limits to evaluate different Harvest Control Rules (HCRs) to achieve the intended probabilities of exceeding the limits and targets in the best possible manner.

Limits and targets for ICES stocks were evaluated using FAO criteria based on MSY $B_{trigger}$ versus $B_{MSY}$ estimated from EQSIM. Based on Figure 4.6.2 some of the reference points are too low and not precautionary in a sense that the targets are sometimes around or less than $0.1B_0$. $B_{trigger}$ is the target, and can be as low as 0.1 $B_0$, where those are instead limits in tunas RFMO's. $B_{lim}/B_0$ is also too low for >50% of the stocks. In addition, if we manage these based on the estimated EQSIM targets the stock status in ICES appears to be a lot worse using the FAO criteria of maximally sustainably fished between 0.8-1.2 $B_{MSY}$ (Figure 3.9).

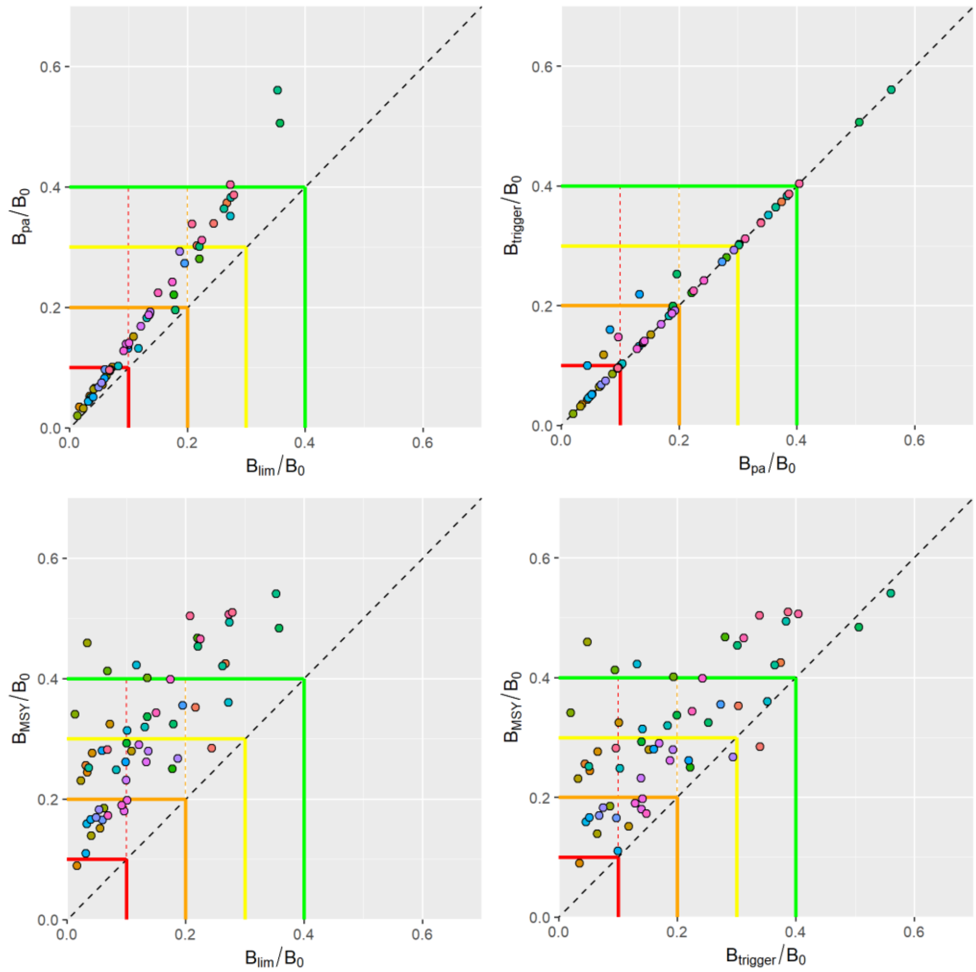


**Figure 3.9. Scatter plots of reference points used in ICES relative to $B_0$ as derived from EQSIM (see Section 2.3).**

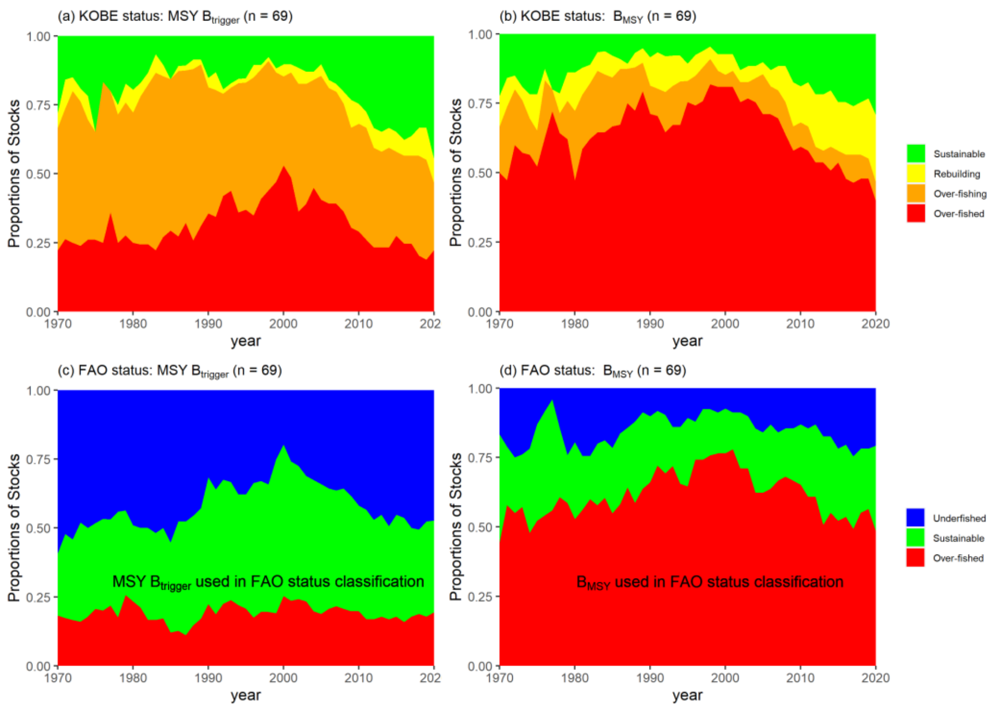


**Figure 3.10. Kobe (upper panels) and FAO (lower panels) status plots of ICES stocks using different limits and targets.**

## 3.4 Reference Points in the U.S. Fishery Management System

*by Richard Methot*

The fishery management system in the U.S. is built upon the Maximum Sustainable Yield (MSY) concept as prescribed in the Magnuson-Stevens Fishery Conservation and Management Act (1976). In particular, the Act defines National Standard 1 as "Conservation and management measures shall prevent overfishing while achieving, on a continuing basis, the optimum yield from each fishery for the United States fishing industry". MSY, as a long-term average concept, is the upper limit on the long-term Optimum Yield (OY) which takes into account ecosystem, economic and other considerations, including uncertainty. Subsequent revisions of the Act and associated guidelines for NS1 have created a well-defined system with strong accountability. Eight Fishery Management Councils have created 46 Fishery Management Plans to cover the 460 managed stocks and stock complexes under federal jurisdiction.

The system includes a Maximum Fishing Mortality Threshold (MFMT), which is typically set at $F_{MSY}$ or a SPR-based proxy for $F_{MSY}$. The level of this proxy is defined in the FMPs and typically ranges from $FSPR_{30\%}$ to $FSPR_{45\%}$. An overfishing limit (OFL) is the level of catch corresponding to application of MFMT to the current stock abundance. Catch>OFL or F>MFMT invokes a condition of overfishing beyond which curtailment or cessation of fishing for the remainder of the fishing year is often required. Probability of overfishing is reduced below 50% by requiring an ABC (Acceptable Biological Catch) control rule below the OFL according to the degree of scientific uncertainty in the OFL. This is termed the P* approach for data-rich assessments, then tiered alternatives for data-limited assessments. These ABC control rules typically have an inflection such that the target F is reduced as the stock abundance declines below $B_{MSY}$. The Councils specify ABC control rules in their FMPs and their Scientific and Statistical Committees apply the ABC control rules to recommend ABC. The Councils are then bound to set the Annual Catch Limit (ACL; e.g. quota) not to exceed the ABC. So, the overall system has ACL <= ABC < OFL.

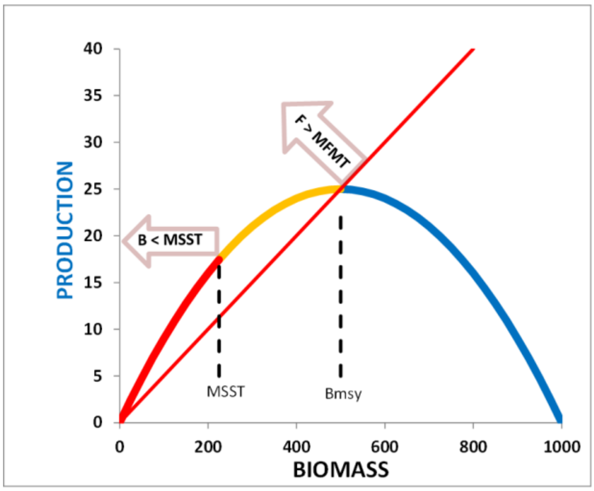


**Figure. 3.11. Reference points in the U.S. management system. Parabolic arc is the production curve. Slope of the diagonal red line is the level of F that is the Maximum Fishing Mortality Threshold (MFMT) above which overfishing occurs. Value of the red line is the amount of catch that would be overfishing. Right-hand vertical dashed black line is the level of biomass associated with maximum productivity, termed $B_{MSY}$. Left-hand vertical dashed black line is level of biomass termed the Minimum Stock Size Threshold (MSST) below which a stock is overfished and for which a rebuilding plan is required.**

The system includes a Minimum Stock Size Threshold (MSST) below which the stock is considered to be overfished and invokes requirement for a rebuilding plan. The MSST is explicitly linked to the $B_{MSY}$, or $B_{MSY\ proxy}$, which is the biomass of the stock associated with $F_{MSY}$. Typically, MSST is set at 0.5*$B_{MSY}$, but some FMPs set MSST closer to $B_{MSY}$. The system requires reporting of situations in which a stock is expected to become overfished in two years, but this is difficult to implement routinely, partly because not all stocks are assessed every year. The F associated with MSST is not routinely calculated or reported, but there may be merit in doing so in order to help identify situations in which a stock is approaching an overfished condition. A semantic and communication challenge with the U.S. system is that the overfished threshold does not correspond to the overfishing threshold. A stock can be experiencing overfishing, yet be above the overfished threshold and be increasing in abundance.

When a stock is determined to be overfished, the system requires a plan, with exceptions, to rebuild the stock to $B_{MSY}$ within 10 years, or the minimum time to rebuild with no fishing plus one mean unfished generation time (with alternatives), whichever is shorter. This has the challenging effect of a pinch point for stocks that can rebuild in just under 10 years with F=0. The rebuilding plan is effectively a temporary replacement for the ABC control rule until the stock is rebuilt.

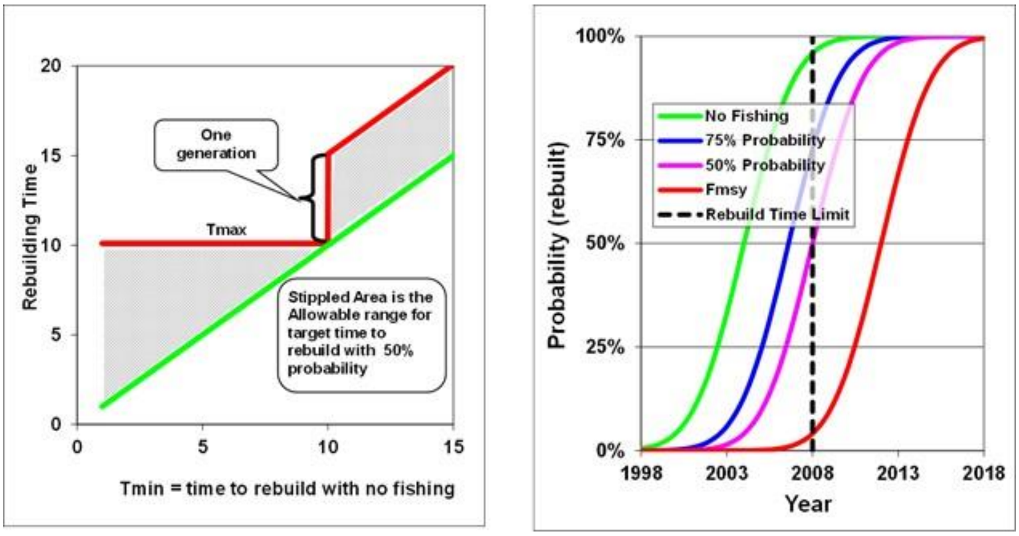


**Figure 3.12. Left-hand panel shows the quantities associated with calculation of $T_{max}$, the maximum allowable time for a 50% probability of rebuilding a stock to $B_{MSY}$. In the right-hand panel, the vertical dashed line at 2008 represents the year in which time to rebuild would be at $T_{max}$. Four cumulative distribution curves show the probability distribution for time to rebuilding given various levels of fishing during the rebuilding period.**

The U.S. system often results in managing the weak species in multi-species fisheries. The optimum level of F for a major target species in such a fishery may cause an associated F on other species to exceed that weaker species' MFMT. This can lead to forgone yield for the major target species to prevent exceedance of the MFMT for weaker species. This is required when the weaker species is on a rebuilding plan. In other situations, the guidelines include a mixed stock exception in which a multi-species analysis could demonstrate the higher regional benefits that could be attained by allowing the weaker species to have an F up to the level of F associated with MSST.

The greatest challenge though is non-stationarity in the productivity and spatial distribution of species. Regime shifts and long-term climate drift in these characteristics, which form the paradigm for reference points, create great difficulty for the definition and updating of reference points to reflect prevailing environmental conditions.

## 3.5 Reference Points in the Canadian Fishery Management System

Part of the recent work by WKRPChange (ICES 2021b) involved reviewing the basis of the ICES reference points, and contrasting the ICES procedures with those in the USA and Canada in the context of providing guidance for future reference point estimation within ICES. Similarly, WKREBUILD (ICES 2020a) reviewed the details of the Canadian reference point system and provided a comparison of the ICES, Canadian, US, and NAFO approaches with focus on defining limit and trigger reference points. Here, a number of key elements of the Canadian reference point system are revisited in the context of specific guidelines to estimate reference points as part of Canada's harvest strategy.

Fisheries and Oceans Canada's (DFO) Fishery Decision-Making Framework Incorporating the Precautionary Approach (PA Policy; DFO, 2009) is a policy that describes a framework where reference points and harvest decision rules are used to make management decisions to support sustainable fisheries. The PA Policy also provides general guidance on determining reference points. In general, reference points should be based on the best information available on the stock biology and fishery while taking into account the limitations of the available data. Specific values for individual stock harvest strategies are to be provided by the scientific stock assessment process.

Uncertainty and risk associated with the estimation of stock status, reference points and in making and implementing management decisions are explicitly recognized in Canada's harvest rate strategy. The guidelines stipulate that uncertainty should be incorporated in the calculation of stock status and biological reference points. Risk is expressed by the identification and position of reference points, the changing severity of management actions that are chosen as stock status changes and the tolerance for stock declines.

Ideally, the maximum acceptable harvest removal reference point is determined analytically as the best estimate of $F_{MSY}$ from the stock assessment model. The advised fishing mortality ($F_{trg}$) can be at or below $F_{MSY}$, but must not exceed it, i.e. $F_{trg} \leq F_{MSY}$. $F_{trg}$ may be set smaller than $F_{MSY}$ by factoring in the impact on other stocks ecosystem considerations and precaution in light of uncertainty.

The stock status zones are defined as the Limit Reference Point (LRP) at the Critical-Cautious zone boundary, and an Upper Stock Reference Point (USR) at the Cautious-Healthy zone boundary and the Removal Reference for each of the three zones (Figure 3.13). The LRP represents the stock status below which serious harm is occurring to the stock based on biological criteria and established by Science through a peer reviewed process. There are several approaches for calculating the LRP in use, which may need refinement over time. The USR is the threshold point below which $F_{trg}$ must be reduced. In practice, the position of the USR is guided by the Target Reference Point (TRP) that corresponds to the estimate for $B_{MSY}$.

In absence of a pre-agreed harvest rule developed in the context of the precautionary approach, DFO (2019; Appendix 1b) provides provisional guidance that is indicative of how the range of each stock status zone should be characterized and where an LRP and USR should be situated along the spectrum of a stock's potential status. The following reference points are emerging from review and meta-analyses of experience with a wide variety of fish stocks. These are in line with practices and standards used internationally, such as New Zealand and USA, informed by scientific review and meta-analyses, and consistent with the language found in various international agreements (Shelton and Sinclair, 2008).

***Critical Zone***: The stock is considered to be in "the critical zone" if the mature biomass, or its index, is less than or equal to 40% of $B_{MSY}$ ($B \leq 0.4\ B_{MSY}$), with $F_{trg} = 0$.

***Cautious Zone:*** The stock is considered to be in the "cautious zone" if the biomass, or its index, is higher than 40% of $B_{MSY}$ but lower than 80% of $B_{MSY}$ (0.4 $B_{MSY} < B <$ 0.8 $B_{MSY}$). $F_{trg}$ is reduced linearly reduced the between URP and LRP, such that $F_{trg} < F_{MSY}(B - 0.4B_{MSY})/(0.8B_{MSY} - 0.4B_{MSY})$.

***Health Zone:*** The stock is considered to be "healthy" if the biomass, or its index, is higher than 80% of $B_{MSY}$ (B > 0.8 $B_{MSY}$), with $F_{trg} \leq F_{MSY}$.

Within the ICES reference point system, these guidelines would therefore translate into setting $B_{lim}$ = 0.4$B_{MSY}$ and $MSY_{Btrigger}$ = 0.8 $B_{MSY}$, where $B_{MSY}$ is the expected biomass corresponding to $F_{MSY}$.

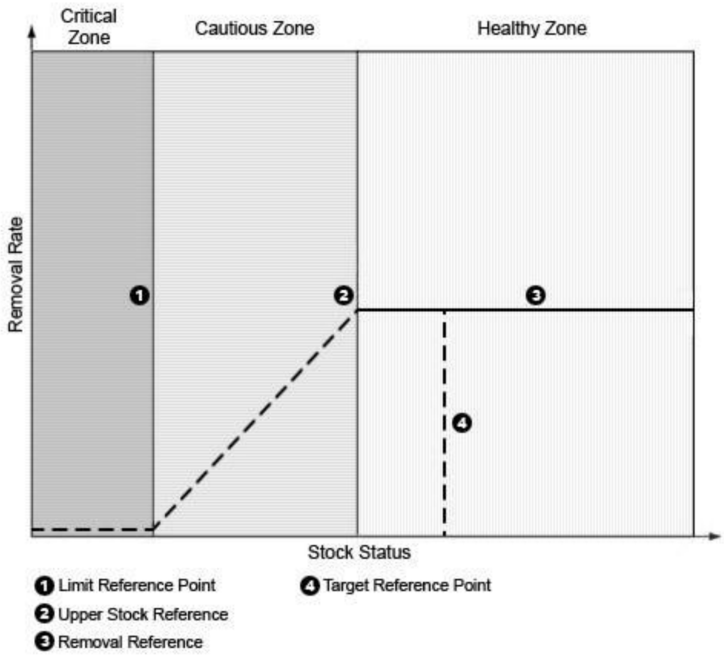


**Figure 3.13. Illustration of the Canadian reference point system, indicating the position of the Limit and Upper Reference Point relative to the Target Reference Point (Source: DFO, 2009).**

## 3.6 Reference Points in the New Zealand Management System

The Harvest Strategy Standard for New Zealand Fisheries (New Zealand Ministry of Fisheries 2008) provides perhaps one of the most unambiguous and transparent frameworks for setting fishery and stock targets and limits and associated fisheries management measures. The framework is based on an operationally explicit definition of sustainability and specifies probabilities for achieving the targets and not breaching limits, with clear commitment to attain long-term biomass levels above $B_{MSY}$ by a probability of more than 50%. The metrics specified are to be treated as defaults and should be applied in most situations, whereas departures must first be justified in terms of the particular circumstances. Acknowledging this, Shelton and Sinclair (2008) proposed to draw on the New Zealand Standard to sharpen the definitions of sustainable harvest strategy implementation in the Canadian framework.

The New Zealand Standard consists of three core elements: (1) a specified biomass target about which a fishery or stock should fluctuate; (2) a "soft" biomass limit reference point that triggers a requirement for a formal, time-constrained rebuilding plan (c.f. ICES 2020; WKREBUILD); and

(3) a hard limit below which fisheries should be considered for closure. These target and limit reference points are specified on the basis of so called "MSY-compatible reference points", which encompass direct estimates of $B_{MSY}$, $F_{MSY}$ and MSY, as well as their analytical and conceptual proxies. The defaults for the biomass targets and limits are specified as follows:

***Biomass target**:* The default target is $B_{MSY}$ or better (more conservative) with at least a 50% probability of achieving the target.

***Soft-limit:*** 0.5 $B_{MSY}$ or 0.2 $B_0$, whichever is higher. The soft-limit is considered breached and the stock classified as depleted if there is a more than 50% probability that the biomass falls below the soft limit.

***Hard-limit:*** 0.2 $B_{MSY}$ or 0.1 $B_0$, whichever is higher. The hard-limit is considered breached and stock classified as collapsed. if there is more than 50% that the biomass is below the soft limit.

Overfishing will be deemed to be occurring if $F_{MSY}$ (or its proxy) is exceeded on average (3-5 years), so that $F_{MSY}$ is the maximum acceptable target fishing mortality. If biomass falls below the trigger point located between the biomass target and the soft-limit fishing mortality is reduced linearly to keep the stock close to the target and away from soft-limit.

Harvest strategies based on Management Strategy Evaluations (MSEs) are advocated and fully-compatible with the minimum requirements of the Harvest Strategy Standard. The default performance criteria for MSEs are therefore specified to ensure that: (1) the probability of achieving the $B_{MSY}$ target or better is at least 50%, (2) the probability of breaching the soft limit does not exceed 10%, (3) and the probability of breaching the hard limit does not exceed 2%. To improve computational efficiency of the MSE, amalgamating the soft and hard-limit into a single criterion of "no more than a 5% probability of breaching the soft limit" is generally accepted.

Detailed Guidance on methods for calculating the reference points, including their proxies, is provided in the Operational Guidelines (New Zealand Ministry of Fisheries 2011). In practice, very few stocks are managed using direct estimates of $B_{MSY}$ (Punt *et al.*, 2014). Reasons given in Operational Guidelines for when direct estimates of $F_{MSY}$ or $B_{MSY}$ are unreliable include strong assumptions about the shape or the steepness (e.g. fixing) of stock-recruitment relationships. In such cases, analytical proxies are considered more reliable and credible. Guidelines for default proxies for $B_{MSY}$ as a ratio to $B_0$ and $F_{MSY}$ based on $F_{\%SPR}$ (Table 3.2), which were developed based on a comprehensive review of the fisheries science and management literature, supported by model age-structured modelling of the relationships (Mace, 1994). The default proxies are based on the productivity of the stock, which are assigned on the basis of biological characteristics, including the intrinsic rate of population increase *r*, maximum age, age at first maturity and generation time. Although default targets may be lower, it is proving difficult to justify MSY targets less than 30-40% $B_0$, also with respect to the soft-limit at 20% $B_0$.

**Table 3.2. Recommended default proxies for $B_{MSY}$ (expressed as %B0) and FMSY (expressed as F%SPR levels from spawning biomass per recruit analysis). Source: New Zealand Ministry of Fisheries (2011)**

| **Productivity Level** | **% $B_0$** | **$F_{\%SPR}$** |
|---|---|---|
| High* | 30% $B_0$ | $F_{SPR35}$ |
| Medium | 35% $B_0$ | $F_{SPR40}$ |
| Low | 40% $B_0$ | $F_{SPR45}$ |
| Very low | > 45% $B_0$ | $F_{SPR50}$ |

* Here adjusted from 25% $B_0$ and $F_{SPR30}$, considering the note that it is increasingly difficult to justify MSY-compatible targets less than 30%-40% $B_0$

## 3.7 Limitations and inconsistencies considering international best practice, the precautionary approach and international policies and legislation

According to international best practices, reference points should be based on the best available information on the stock biology, fishery and the ecosystem, while taking into account the limitations of the available data. In the ICES system, this translates into the benchmark process, which consists of periodically (i.e. every 3 to 5 years) revising data, the assessment model and the reference points for each stock. The benchmark process provides the expert knowledge base on the stock's biology and its surrounding ecosystem, undergoes external review and is therefore best suited for specifying reference points for individual stocks that are consistent with the assessment assumptions.

According to international best practices, $F_{MSY}$ is almost invariably set as the maximum acceptable or limit level of fishing mortality to which the stock should be exposed. Due to the inherent difficulties of estimating $F_{MSY}$ reliably and taking into account uncertainty of several of the key parameters describing the stock and fisheries dynamics, $F_{MSY}$ is often replaced by more conservative proxies, of which the widely used $F_{spr\%}$ corresponds to a spawner-biomass per-recruit (SPR) ratio of 30 to 50% of $SPR_0$ and $F_{B\%}$ corresponds to a SSB between 25% and more than 45% of $B_0$. The different fractions of SPR and $B_0$ typically vary according to the biology and ecology of the assessed species. Internationally, $B_{MSY}$ is the target biomass ($B_{trg}$) reference point that corresponds to the SSB at $F_{MSY}$, the $F_{MSY}$ proxy, or to a fraction of $B_0$, which usually range between 25 to 45+% of $B_0$. The benefit of relating the target biomass reference point to $B_0$ instead of $B_{MSY}$, is that it is more robust to assumptions about the SRR and selectivity pattern of the fisheries (Hilborn, 2010). $B_{lim}$, the biomass limit reference point *sensu* ICES, is generally informed either by a fraction of $B_{MSY}$ or $B_0$, with the given fraction being stock and species dependent and chosen by taking into account the biology and resilience of the stock under assessment. Finally, $B_{trigger}$ the operationalized biomass reference point, is set as a fraction of $B_{MSY}$ as in Canada (i.e. 80% of $B_{MSY}$) or equal to $B_{MSY}$ as in the USA.

When analysing the position of the current biomass reference points in the ICES system compared to $B_{MSY}$ and $B_0$ (i.e. the key biological biomass reference points according to international standards) for the 69 category 1 stocks included in the ICES database, there was a very large variation in ratios of $B_{lim}$ to either $B_{MSY}$ or $B_0$ (Figure 3.14). The ratios of $B_{lim}$ to $B_0$ vary from very small fractions of 1.3% of $B_0$ to 38% of $B_0$, with 50% of $B_{lim}$ values located below 10% of $B_0$ (median = 0.1 $B_0$). The extreme ranges for both $B_{pa}$ and MSY $B_{trigger}$ are not much higher (0.02 - 0.56 $B_0$) with the median increased to 14%, reflecting the relationship $B_{pa} = 1.4B_{lim}$. The ratio of $B_{lim}$ to $B_{MSY}$ also spans a wide range from less than 4% to 85.8% (median = 0.39 $B_{MSY}$), and the ratio MSY $B_{trigger}$ to ranges from 5.8% to 120% of $B_{MSY}$ (median = 0.6 $B_{MSY}$). Meaningful location compared to the biology of the species, determines the effectiveness of both MSY $B_{trigger}$ and precautionary reference points $F_{P.05}$. Setting $B_{lim}$ well under 10% of $B_0$ renders $F_{P.05}$ ineffective for most ICES stocks with or without the use of $B_{trigger}$ (Figure 3.15). This is particularly important in the presence of the Allee effect in exploited fish, which was identified to occur when the stock is below 15-25% of $B_0$ (Perälä and Kuparinen 2017; Perälä et al., 2021). Finally, the use of $B_{trigger}$ when estimating $F_{MSY}$, has the effect of increasing both $F_{P.05}$ and $F_{MMY}$, resulting in an $F_{MSY}$ that is on average 12% higher if running EQSIM with harvest rule than without.

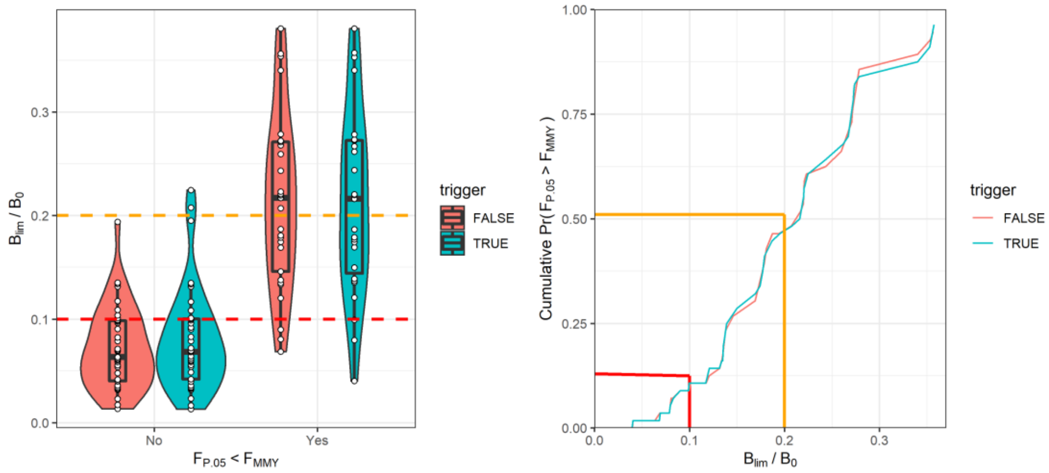


**Figure 3.14. Plots showing (left) comparisons of stocks for which $F_{P.05}$ was larger than $F_{MMY}$ (Yes/No) in relation to the ratio $B_{lim}$ /$B_0$ for EQSIM was run (Section 4.3) with MSY $B_{trigger}$ (TRUE) or without MSY $B_{trigger}$ (FALSE) and (right) the cumulative probability of invoking Precautionary $F_{P0.5}$ as a function of $B_{lim}$/$B_0$ (n = 69 Category 1 Stocks *sensu* ICES of which 55% did not invoke $F_{P.05}$)**

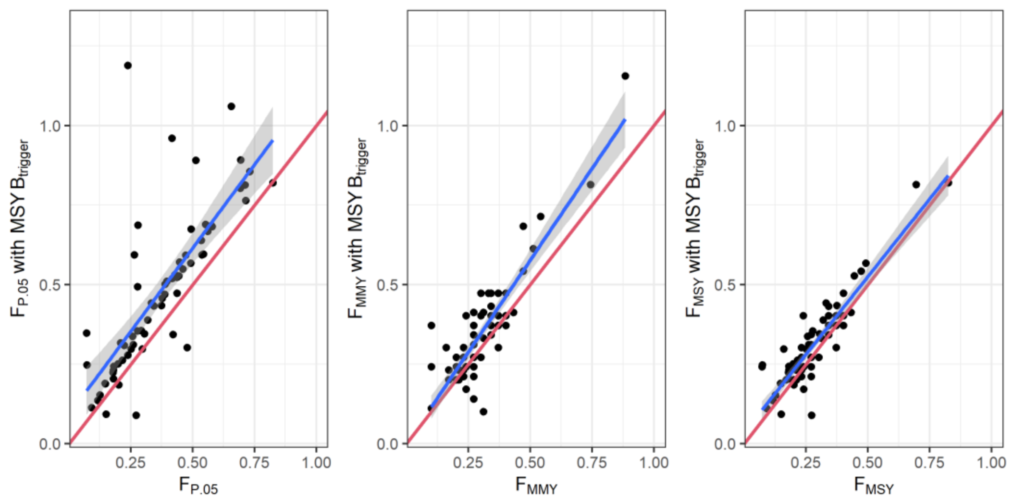


**Figure 3.15. Effect of using or not MSY $B_{trigger}$ on the EQSIM estimates of $F_{P.05}$, $F_{MMY}$, and $F_{MSY}$ (Section 4.3).**

# 4 Evaluate the robustness, consistency and plausibility of limit and target reference points in relation to current ICES Advice Rule in comparison to alternative approaches.

## 4.1 The uncertainty of $B_{lim}$ and some input to the existing ICES guidelines on reference point estimation

*by Mikael van Deurs*

In a study published in 2021, we used data from 51 pelagic fish stocks and evaluated the robustness of type-1 and type-2 methods for estimating $B_{lim}$ (ICES guidelines) (van Deurs *et al.,* 2021). Using a combination of data simulations and data from 51 small-bodied pelagic fish stocks, we analyzed the sensitivity of $B_{lim}$ to (a) the choice of method, (b) time-series length and (c) stock development (e.g. rebuilding or declining). It was demonstrated that $B_{lim}$ is associated with considerable uncertainty. Furthermore, $B_{lim}$ and the level of uncertainty depended on choice of method, time-series length and stock development trends. Lastly, we propose a simple bootstrap approach for quantifying $B_{lim}$ uncertainty. The nonparametric methods defining "large" recruitment as greater than the 50th percentile or the 80th percentile (type-1) provided relatively precise estimates of the true $B_{lim}$ (i.e. narrow confidence intervals). However, these methods were prone to bias (i.e. deviated from "true" $B_{lim}$ in simulation study) and were sensitive to time-series length. Most often $B_{lim}$ declined with increasing time-series length, although this depended on stock development (increasing or decreasing trend in SSB over time). When the variation around the SR relationship was low, the hockey stick approach (type-2) was much less biased and more precise regardless of time-series length and stock development and hence should be the preferred option. However, when variability around the SR relationship is large, it may be necessary to also consider $B_{lim}$ from other methods in addition to the hockey stick method. In conclusion, the choice of method was the single most influential factor in estimating $B_{lim}$ for the four stocks. Without clear guidelines for when to choose each of the methods, this introduces a hidden subjectivity and lack of transparency in stock management. However, although some generic patterns were identifiable, analysis on a stock-by-stock basis (using real stock data) showed large variation among stocks and many stocks did not conform to the general patterns described above. Besides what is presented in the paper described above, the present study on $B_{lim}$ estimation also gave rise to a set of suggestions for changes to the existing ICES guidelines on $B_{lim}$ (ICES 2021a):

- In case of type-1: It should be specified what "good recruitment" is. A suggestion could be that any recruitment higher than the 70% lowest recruitments is a good recruitment, since this is roughly comparable to the type-2 approach in those cases where type-2 results in a $B_{lim}$ below half of the 95% upper conf. bound of SSB (Figure 4.1a). Furthermore, type-1 should mainly be attempted in those cases where a realistic SR relationship (incl. a segmented regression or hockey stick) cannot be fitted (and before moving on to type-3+).
- Decide on what method to use for estimating break point in hockey stick (grid-search method or the bent hyperbola methods); or don´t choose type-2 unless both methods give the same result. The two methods sometimes yield very different results (Figure 4.1b).

The grid search approach is recommended by Barrowman and Myers (2000), but has a long run time in R if the grid-steps are small.

- Use a "boot-strap" approach to estimate $B_{lim}$ uncertainty. See example in Figure 4.1c of a boot-strapped $B_{lim}$ distributions (type-2). Provide the uncertainty as an indicator pf robustness.
- In case of type-1: Instead of choosing lowest SSB yielding good recruitment (i.e. $B_{lim}$ is hinged on one value), then base it on the geom. mean of the $B_{lim}$ distribution derived from the boot-strapping (Figure 4.1d) (perhaps consider taking the geom. mean after removing outliers outside the 90% confidence interval). This approach will of course be less functional for some special cases, such as extremely spasmodic stocks with very few "successful" recruitment years (and the rest being merely recruitment failures), in which case a different procedure should be applied. If we want to take into account estimation uncertainty of recruitment and SSB (from the assessment), this could easily be included in the boot-strap procedure (i.e. when sampling a specific recruitment-SSB pair in the time-series, a value is drawn from a distribution specified by the estimation uncertainties of this recruitment-SSB pair).
- Develop an R-tool to be used in the working groups for consistency on the bootstrap procedure.
- Benchmark working groups should be requested to clearly report the steps they took when estimating $B_{lim}$. Perhaps developing a decision tree, could be useful (similar to what was done in relation to judgement of retrospective bias at the WKFORBIAS workshop in Woods Hole in 2019).

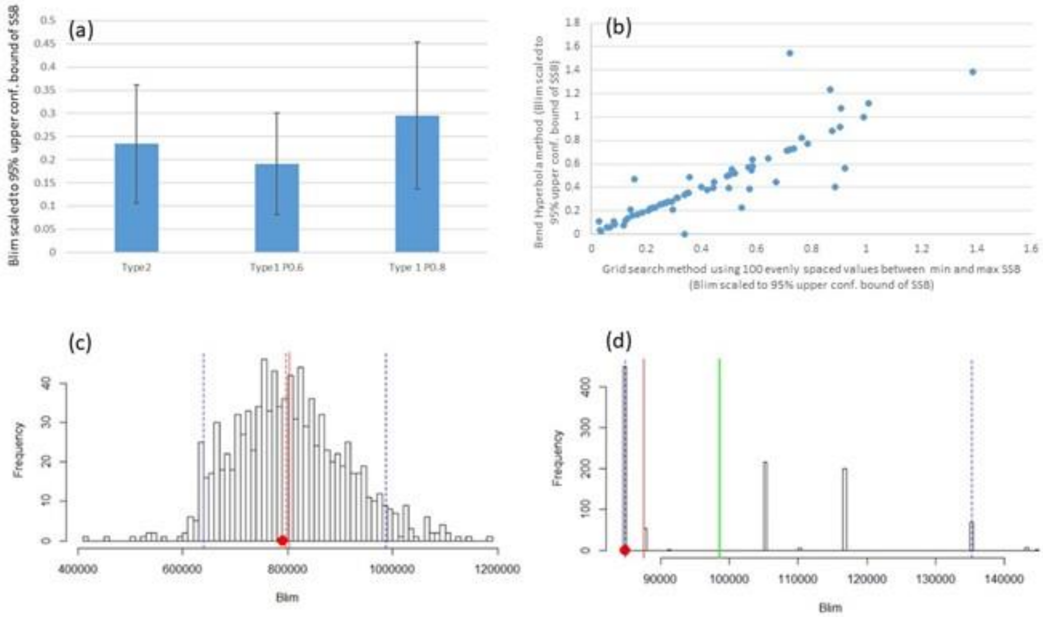


**Figure 4.1. (a) Average $B_{lim}$ (and standard deviation) from using type-1 and type-2 approach (good recruitment = any recruitment larger than the 60% (P0.6) and 80% (P0.8) lowest recruitment values, respectively). The figure is based on stocks included in van Deurs *et al.* (2021) (each data point represents a stock), after excluding those stocks where type-2 approach produce $B_{lim}$ lower than half of the 95% upper conf. bound of SSB). Grid search method was used to identify hockey-stick break point. (b) Comparison of two methods for estimating the break point of the hockey stick when choosing type-2. Based on stocks included in van Deurs *et al.* (2021). (c) Example (North Sea herring) of a boot-strapped $B_{lim}$ (type-2) distribution (1000 boot-strap samples with replacement). Red dot is the $B_{lim}$ estimated using all data. Blue dashed lines are the 5% and 95% conf. bounds of the distribution. Dashed red line is the median of the boot strapped distribution and solid red line is the average. (d) Example (North Sea cod) of a boot-strapped $B_{lim}$ distribution (type-1). Green solid line is the geom. mean of the boot-strapped distribution.**

## 4.2 The Management Strategy Evaluation and Reference Points

*by José De Oliveira*

It is not uncommon for requesters of advice to ask for a Management Strategy Evaluation (MSE) of a set of harvest control rules, and for reference points to be estimated as part of this evaluation. A recent example is the MSE for mackerel (ICES, 2020a). What may happen when doing this, however, is that reference points that are derived from an MSE conflict with reference points derived through a more generic framework (EQSIM, the general framework used within ICES to derive reference points for Category 1 and 2 stocks; ICES, 2021), even though the basis for both frameworks is the same accepted benchmark assessment for a particular stock. This is not surprising given the simulation basis and assumptions underlying the two frameworks (e.g. Table 4.1).

The example used for this presentation was the EU-Norway request for MSEs to be conducted for five jointly-managed North Sea stocks, performed by ICES through WKNSMSE (ICES, 2019). One of the outcomes of this workshop was that fishing mortality reference points inferred from the MSEs conflicted with existing reference points derived by the EQSIM framework. Figure 4.2 illustrates the problem for North Sea cod. The evaluation shown was for a hockey-stick-type HCR with a breakpoint at $F_{trg}$ and $B_{trigger}$ and no stability mechanisms (such as TAC constraints, or banking and borrowing). When setting $B_{trigger}$ to MSY $B_{trigger}$ for North Sea cod (which was 150 000 t at the time the MSE was conducted), the F value that was just within the precautionary zone (here $F_{trg}$ = 0.37) would fulfil the definition of $F_{P.05}$ (the fishing mortality that, when used in conjunction with the ICES advice rule – i.e. with $B_{trigger}$ set to MSY $B_{trigger}$ – would lead to 5% probability of SSB being below $B_{lim}$). At the time of the MSE, $F_{MSY}$ for North Sea cod was 0.31 with $F_{MSY\text{-}lower}$ and $F_{MSY\text{-}upper}$ values of 0.198 and 0.46, respectively (Table 4.2). This implies that $F_{MSY\text{-}upper}$ would no longer be considered precautionary in the context of the MSE.

This exercise was repeated within WKNSMSE for all five stocks under consideration (ICES, 2019). There were conflicts with reference points from the two frameworks in all cases but one (haddock). This raises the question about which framework is most suitable for deriving reference points: the more generic EQSIM framework, or the MSE framework that can be better tailored to the uncertainties relevant for a particular stock. There is a strong argument, put forward by WKGMSE3 (ICES, 2020b), for the MSE framework to be used to derive the reference points when an MSE is being conducted for a stock, and this workshop even made a proposal for how this could be done (see section 2.3 of ICES, 2020b). This is because an MSE accounts for uncertainty in a more comprehensive manner than EQSIM, which is likely to be important when estimating metrics such as $F_{P.05}$ (which come from the tails of a distribution). Furthermore, WKMSEMAC (ICES, 2020a) also supported the use of an MSE framework to derive reference points when one is conducted, arguing that an MSE framework:

- Includes the actual assessment and forecast process and can therefore more appropriately handle the related uncertainties;
- Is a consistent approach with that used for the evaluation of the long-term management strategy?

A discussion about the use of MSEs to derive reference points (when MSEs are conducted) will be taken up during the second WKREF meeting (WKREF2) to be held in early 2022.

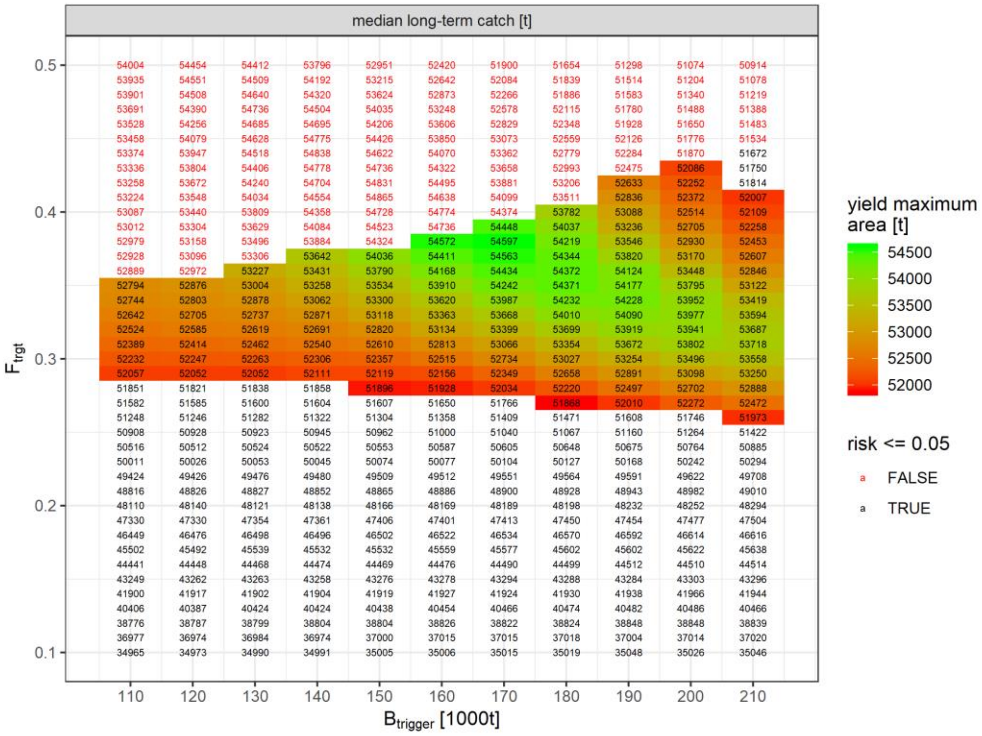


**Figure 4.2. Median long-term catch for different combinations of the control parameters $F_{trg}$ and $B_{trigger}$. $F_{trg}$-$B_{trigger}$ combinations in red text are not precautionary (more than 5% probability of SSB<$B_{lim}$), while those in black text are precautionary (no more than 5% probability of SSB<$B_{lim}$), and within these, those cells having the highest median catches shaded (in "heat" colouration indicated by the legend). The cell with the black border is $F_{P.05}$ with the ICES advice rule ($B_{trigger}$=MSY $B_{trigger}$). Reproduced from ICES (2019).**

**Table 4.1. Difference between EQSIM and the MSE framework used for WKMSEMAC (reproduced from ICES, 2020a).**

| | EqSim | MSE framework |
|---|---|---|
| Assessment/advice Error | 2 parameter ($F_{cv}$, $F_{phi}$) function derived from historic performance of assessment and forecast, adding auto-correlated (but unbiased) noise to the target fishing mortality. | Full feedback approach. Includes an assessment with consistent bias in estimates of SSB/FBar such that the realized F is always lower than the target fishing mortality. |
| Recruitment residuals | Process error generated around SR. Identical starting population for all iterations. Autocorrelation in residuals estimated and included in predicted residuals. | SR models assigned randomly to the 1000 populations derived from the variance-covariance matrix. Auto-correlated random deviations derived from ARIMA fit to log residuals for each iteration. |
| Simulation/Stat period | 200 years/ 50 years | 40 years/ 5 years |

**Table 4.2. The fishing mortality reference point $F_{P.05}$ inferred from the MSEs conducted for five jointly-managed North Sea stocks (ICES, 2019), compared to $F_{MSY}$ and associated ranges derived using EQSIM (values from published advice in 2019). Shaded cells are those reference points derived from EQSIM that would not be considered precautionary under the MSE.**

| | MSE | EQSim | | |
|---|---|---|---|---|
| Stock | $F_{P.05}$* | $F_{msy\text{-}lower}$ | $F_{MSY}$ | $F_{MSY\text{-}upper}$ |
| Cod | 0.37 | 0.198 | 0.31 | 0.46 |
| Haddock | 0.23-0.26 | 0.167 | 0.194 | 0.194 |
| Whiting | 0.10-0.11 | 0.158 | 0.172 | 0.172 |
| Saithe | 0.37-0.42 | 0.210 | 0.363 | 0.536 |
| autumn-spawning herring | 0.22 | N/A | 0.26 | N/A |

* The reason for a range and not a single value in some cases is because only partial $F_{trg}$-$B_{trigger}$ grids were possible in most cases (unlike the full grid shown in Figure 1), but the $F_{P.05}$ value could still be inferred to be within this range.

## 4.3 The Evaluation of the robustness of alternative fisheries reference point systems

*by Henning Winker, Massimiliano Cardinale, Iago Mosqueira, Laurence Kell, Rishi Sharma, Christoph Konrad and Michaël Gras*

Central to fisheries advice is the reference point system, which, in a first instance, may be used to classify and communicate current status of the resource, but is ultimately designed to determine, e.g., the total allowable catch (TAC) in managed fisheries. The stock assessment model is often considered the starting point for management advice. However, the process starts in reality with imperfect observations that are typically associated with large observation and systematic sampling errors. The assessment model itself relies on many assumptions about the model structure in the form of the underlying deterministic relationships (e.g. the stock-recruitment function) and key population parameters (e.g. natural mortality M). All of these contribute to the uncertainty associated with the stock assessment output (Patterson *et al.*, 2001), where uncertainty can be seen as a plausible range of differences between the model outcomes and reality.

This analysis seeks to evaluate the robustness of the current ICES reference point system against alternative approaches. In statistics, the term robust refers to a test or model that provides correct inference despite its assumptions being violated, whereas a robust system in engineering is one that remains to function correctly in presence of uncertainty (Radatz *et al.*, 1990). In the context of fisheries advice both meanings are interrelated and highly relevant (Kell *et al.*, 2016). Evaluating the robustness of a reference point system therefore requires testing if it can also produce desired outcomes in situations where the reality (operating model) differs in assumptions from reference point estimators. To do this, we condition operating models based on the ICES 2021 age-structured stock assessment outputs for 69 stocks Category 1 stock assessment outputs with available $F_{MSY}$ estimates for 2021. The conventional ICES "hockey-stick" ICES advice rule was parameterized based on the benchmarks according to the 2021 advice and compared to alternative parameterisations based on five generic reference point estimators. The performance evaluation was conducted using ICES "short-cut" approach to MSE framework (ICES 2020b; WKMSE3), and implemented in the mse package (Mosqueira and Jardim, 2019) in FLR (Kell *et al.*, 2007).

## A Short-Cut MSE Approach to robustness evaluations

The key difference that distinguishes an MSE approach from a simple stochastic risk simulation, such as EQSIM, is the simulated feedback control loop between the implementation system and the operating model (Punt *et al.*, 2017), where the implementation system translates the assessment outcome via a harvest control rule (HRC) into a management quantity, such as TAC advice. In ICES, the implementation system of the harvest control rule is based on the assumption that advice is given for year *y*+1 based on an assessment completed in year *y*, which is typically fitted to data up until last data year y-1 (ICES, 2020b). Therefore, assumptions about catch and population processes during the intermediate assessment year y require a short-term forecast and a prediction of the catch according to the HCR in the management year y+1 (Mildenberg *et al.*, 2021). The advantages of a short-cut MSE for testing the robustness of reference point systems across a large number of stocks under the same conditions are: (1) the easy implementation using within the unified 'FLR' framework using latest 'mse' version associated with 'FLasher', (2) reduced computation time, (3) minimal data requirements in the form of standard assessment outputs (FLStock), and (4) the incorporation of the lag effect between data, assessment and management implementation.

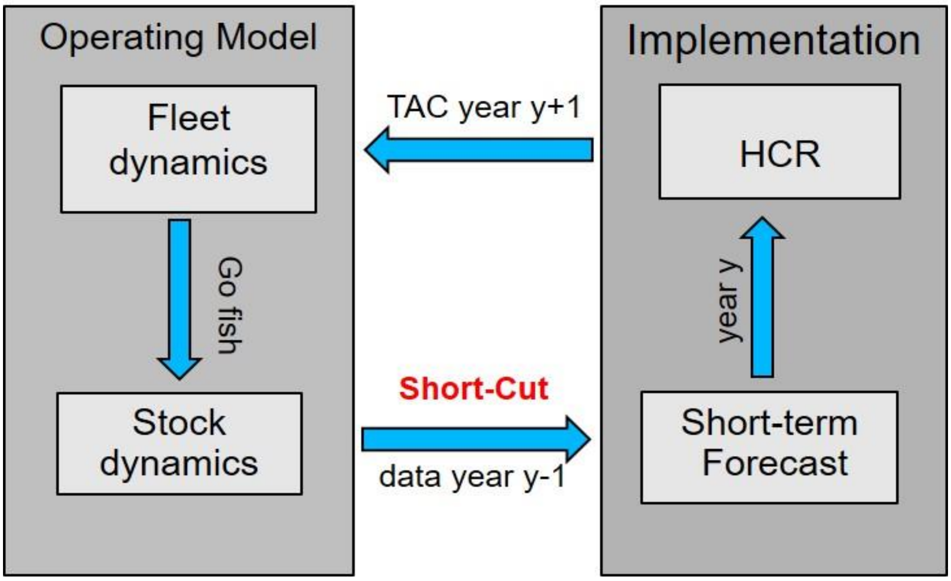


**Figure 4.3. Schematic illustrating the key processes of the short-cut approach to MSE, showing the Operating model that simulates the fishery and stock dynamics on the left and Implementation System including the short-term forecast on the right. The short-cut denotes the omission of the estimation (stock assessment) model that updates to new observations (with estimation error) in a conventional MSE implementations with full feedback control loop.**

## Conditioning of Operating models

The OMs for all 69 stocks were implemented as single sex and single fleet models with an annual time step. Future projections were run over 60 years (i.e. 2021-2080) with 250 iterations and based on the 3-years average of the most recent data years for weight-at-age ($w_a$), maturity-at-age ($mat_a$), natural mortality-at-age ($M_a$) and the $F_a$ pattern determining the selectivity-at-age ($s_a$) This choice was made to account for non-stationary processes in these quantities (Section 6.1).

Applying the same stock-recruitment relationships (SRRs) that were used in the current advice to condition the OMs is challenging to replicate, due a lack of clear documentation (Section 4.5). More importantly, this would implement effectively a "self-test" of the ICES advice rule rather than a robustness evaluation, considering that testing robustness builds on the premises that the advice outcome is robust to key modelling assumptions being violated. For the robustness testing, a generic Beverton-Holt model (BH-SRR) was assumed for all stocks. The recruitment deviation is assumed to be associated with a first-order autocorrelation (AR1) process and a function of recruitment standard deviation $\sigma_r$ and the AR1 coefficient $\rho$ (Johnson *et al.*, 2016). To ensure an objective and unified approach that represents the wide range of life histories, species-specific predictive distributions for steepness s were and expected means for $\sigma_r$ and $\rho$ were sourced from the hierarchical taxonomic FishLife model to fit a BH SRR to the S-R data and generate the recruitment deviations, respectively (Thorson, 2020; https://github.com/James-Thorson/FishLife).

The parameters of stock-recruit curves are notoriously difficult to estimate, and often little inference can be made from a single stock-recruit fit, but meta-analysis and the use of distributions as a Bayesian prior can provide a useful starting point from which meaningful updates could occur. This approach of using prior information to condition the SRR to the S-R data, is consistent with discussions and suggestions for future work in WKMSYREF2 (ICES 2014). Instead of assuming that nothing is known when fitting the SRR, other than the information that is contained in the stock data alone, this approach assumes that at least within taxonomic groupings (family, species) information from one stock can provide some useful prior information for another (Myers *et al.,* 1999; Thorson, 2020). For stocks with few years of S-R data, or where the observations appear uninformative, priors can assist in making less spurious inference about the SRR, whereas if the S-R data are informative, the priors are effectively updated by the data.

The Beverton-Holt SRRs were fitted to S-R data using the R package FLSRTMB (https://github.com/flr/FLSRTMB), which implements a re-parameterised Beverton and Holt SRR as a function of steepness $s$ and annual unfished spawning biomass per-recruit $SPR_0$ to accommodate the integration of priors for s (Thorson, 2020). A notable difference to the conventional parameterization is that $SPR_{0_y}$ is treated as non-stationary, being a function of annual quantities of $W_{a,y}$, $Mat_{a,y}$ and $M_{a,y}$. By way of using time-varying $SPR_{0,y}$, it also takes into consideration the recent criticism by Miller and Brooks (2021) that specifying a set biological parameters to define a single time-invariant $SPR_0$ can be highly sensitive to reference estimation when using steepness values from meta-analysis (See Appendix I for details).

## Current ICES Advice Rule in comparison to alternative approaches

To facilitate comparability of the alternative reference point systems, all considered harvest control rules (HCRs) are kept generic and in the same form of the conventional ICES Advice Rule (ICES 2021a), with the F advice increasing linear from F = 0 to the target fishing mortality ($F_{trg}$) between zero SSB and the biomass trigger $B_{trigger}$.

In the following, we consider six alternative HCRs, including a 'self-test' and the ICES harvest advice rule (Table 5.5.1). The other four HCRs differ in assumptions to various extent from the OM in terms of SRR and associated proxies for $F_{MSY}$, $B_{MSY}$ and the associated $B_{trigger}$ (Table 4.3).

self-test: The self-test is identical to the OM specifications and therefore represents a reference case that assumes knowledge of the "true" reference points. The target fishing mortality $F_{trg}$ corresponds to the 'true' $F_{MSY}$ of the OM and the biomass trigger is set to 0.8 $B_{MSY}$.

*ices*: The ICES harvest advice rule is parameterized using the official 2021 ICES benchmarks for $F_{MSY}$ as target fishing mortality ($F_{trg}$) and MSY $B_{trigger}$ as $B_{trigger}$.

*bevholt*: The HCR for the bevholt approach is based on fitting a BH-SRR to the S-R data, but without any prior information. $F_{trg}$ is set as the directly estimated $F_{MSY}$ and $B_{trigger}$ is set to 0.8 of the corresponding $B_{MSY}$ estimate at equilibrium.

*sb40*: The HCR the sb40 approach is also fitting a BH-SRR to the S-R data without prior information. However, in this case $F_{trg}$ is set to $F_{sb40}$ as proxy for $F_{MSY}$, where $F_{sb40}$ is the fishing mortality that corresponds to $SB_{40}$ = 0.4 $B_0$ at equilibrium. $B_{trigger}$ is set to 0.8 of $SB_{40}$.

*f0.1*: For the *f0.1* approach the $F_{trg}$ is set to $F_{0.1}$ as proxy for $F_{MSY}$. The $B_{trigger}$ is based on $B_{lim}$ from a "precautionary" continuous Hockey-Stick model that constraints $B_{lim}$ to fall within range of $0.1B_0$

< $B_{lim}$ < $0.5B_0$, but otherwise estimates both the break b = $B_{lim}$ and $B_0$ = ab (a = slope) dynamically within these bounds (see Appendix). The $B_{trigger}$ is set according the empirical rule $B_{trigger} = B_{lim}e^{1.645\sigma}$ with σ = 0.3 (Mildenberger et al. 2021, Ralson et al. 2011)

**Table4.3. Settings of the different reference point estimation approaches tested.**

| HCR | SRR | $SPR_{0,y}$ | $F_{target}$ | $B_{target}$ | $MSY_{Btrigger}$ |
|---|---|---|---|---|---|
| *bevholt* | BH (no priors) | 3 years | $F_{MSY}$ | $B_{MSY}$ | 0.8 $B_{MSY}$ |
| *sb40* | BH (no priors) | 3 years | $F_{SB40}$ | $B_{40}$ | 0.8 $SB_{40}$ |
| *f0.1* | HS | 3 years | $F_{0.1}$ | $B_{F0.1}$ | 1.63 $B_{lim}$ |
| *spr40* | None | N/A | $F_{SPR40}$ | $SPR_{40}$ | $SPR_{30}$ |

## Performance Statistics

The last 10 years of the 60-year projection horizon (i.e. 2071-2080) were used for performance evaluations (10 years × 250 iterations). Performances were broadly based on the components of New Zealand and USA reference points (Sections 4.1-4.3) and are generally consistent with recent literature (Hordyk *et al.*, 2019; Fisher *et al*., 2021; Mildenberger *et al*., 2021). The performance statistics are based on the "true" equilibrium quantities $B_{MSY}$, $F_{MSY}$ and MSY of the OMs. These were computed with FLBRP using 3-years average of the most recent data years for weight-at-age ($w_a$), maturity-at-age ($mat_a$), natural mortality-at-age ($M_a$) and the $F_a$ pattern determining the selectivity-at-age ($s_a$) and given the conditioned BH SRR.

The first component entails that the stock should fluctuate around $B_{MSY}$ with a default target of > 50% probably of SSB > $B_{MSY}$. This implies that the highest acceptable fishing mortality would have to be $F_{MSY}$ or less. A "soft" limit was set 0.5 $B_{MSY}$ and assigned a moderately low probably risk threshold of less than 20% that $SSB$ falls below this limit. Three alternative "hard" limit performance statistics evaluated, which should be prevented with a very high probability associated with a risk threshold of less 5% that SBB falls below these limits for reference:

1. < 5% probability of SSB < 0.3BMSY (WKMSYSPICT, 2021)
2. < 5% probability of SSB < 0.1B0
3. < 5% probability of SSB < $max$(0.3$B_{MSY}$, 0.1$B_0$), i.e. whichever is higher.

The probability type "Prob3" was applied to compute the risk for the biomass limits as the maximum of annual probabilities is taken over the last 10 years (2071-2080) and "Prob1" was used for the probability the SSB > $B_{MSY}$ average probability taken across the last 10 years (2071-2080).

In addition, two fisheries performance statistics were included in the form of median ratio of Catch/MSY and Average annual variability in catches (AAV) (e.g. Fisher *et al.*, 2021).

## Robustness test results

The current ICES advice rule was the least robust compared to any other tested reference point approach as judged by the full set performance statistics (Figure 4.4). The ICES $F_{MSY}$ estimate was the highest among all estimators with the median across all stocks and iterations being about 30% higher than in the OM. The median probability of attaining biomass levels less than 20% and median long-term yield was lowest of all tested approaches. This was generally associated with large variation among the 69 stocks (Figure 4.4). The risk of falling below any of three $B_{lim}$ performance statistics was substantially higher than for any other approaches. Despite assuming a "perfect" assessment, including unbiased estimates of SSB and F, the risk thresholds were exceeded in several cases.

The self-test confirms that fishing at the "true" $F_{MSY}$ with perfect knowledge will in most cases fall short of keeping biomass at levels above $B_{MSY}$ with at least 50% probability (Figure 4.4), which can be attributed to the lag between data and advice under stochastic conditions (Mildernberger *et al*., 2021). Similarly, the long-term median catch was slightly below the deterministic value of MSY. Estimating MSY-based reference points directly by fitting the SRRs in the cases of the *ices* and *bevholt* approaches, resulted in higher $F_{MSY}$ estimates compared to approaches that used proxies for $F_{MSY}$. Although the median $F_{MSY}$ median from the *bevholt* approach was on average unbiased with respect to the "true" value of the OM, it was still associated with poor trade-offs in terms of a low probability of attaining $B_{MSY}$, the risk of falling below the $B_{lim}$ and yield loss rather than yield gain. By contrast, the *sb40* estimator produced negatively biased estimates of $F_{MSY}$ (~ -40% bias), yet showed superior performance statistics, including higher long-term yields associated with lower AAVs. The $f_{0.1}$ approach resulted in the highest long-term yield and showed the second best performance in terms of risk, using the "precautionary" Hockey-Stick approach for $B_{lim}$ ($B_{lim}$ > $0.1B_0$) as a basis for setting $B_{trigger}$. Not relying on any assumption about the SRR, the spr40 estimator performed well for most of the performance statistics, but showed an increased risk in some cases of SSB falling below the $B_{lim}$.

The seeming disconnection between the poor performance of the *bevholt* and the overall superior performance of the most conservative *sb40* approach (Figure 4.4), can be best explained by the asymmetric risk associated with underestimating or overestimating $F_{MSY}$ (Hordyk *et al*., 2021). Asymmetric risk describes the phenomenon that one direction of bias for an estimate leads to disproportionately higher risk than if the bias would occur in the other direction. The asymmetric effects of fishing below or above $F_{MSY}$ are well established in the literature (Beverton, 1998; Mace, 2001; Hilborn, 2001; Hordyk, 2021). The consequence of fishing above $F_{MSY}$ is that the biomass will decrease relative to $B_{MSY}$, so that yield levels close to MSY cannot be maintained. To eventually achieve rebuilding requires fishing mortalities lower than $F_{MSY}$. Fishing below $F_{MSY}$ can result in short-term yield loss but in contrast to overshooting $F_{MSY}$ the catch opportunity still exists at higher biomass levels. In comparison to the biomass increase, the long-term loss in yield is relatively small. For example, Beverton (1998) noted that instead of striving for $F_{max}$ "a simple management system based on careful monitoring of fishing effort, biological targets such as $F_{95}$ (i.e. a lower fishing mortality the results in 95% of the maximum yield), and exploitation of a diversity of fish resources may suffice to avert further disaster and hedge against uncertainty." Respero *et al*. (1998) showed that fishing at just 75% $F_{MSY}$ would still yield an average 0.949-0.989 of MSY based on deterministic age-structured models (Mace, 1994) that was parameterized with 600 combination of variations of life history parameters. Hilborn's (2010) concept of 'Pretty Good Yield' is also founded on the principle that fishing near but not at the maximum yield will reduce risk of overfishing and increase robustness to uncertainties with little long-term yield loss. Even fishing under a harvest control rule at $F_{MSY}$ can still be associated with high risk of a stochastic collapse below $0.5B_{MSY}$ as a result of recruitment variability. On the other hand, the risk can be significantly reduced by fishing somewhat below $F_{MSY}$ (Thorson *et al*., 2015). Recently, Hordyk *et al*. (2021) conducted simulations with stock assessment feedback-loop and identified much higher risk to long-term yields and stock biomass when positively biased stock parameter (e.g *M*, steepness) lead to an overoptimistic $F_{MSY}$ than with the equivalent negative bias.

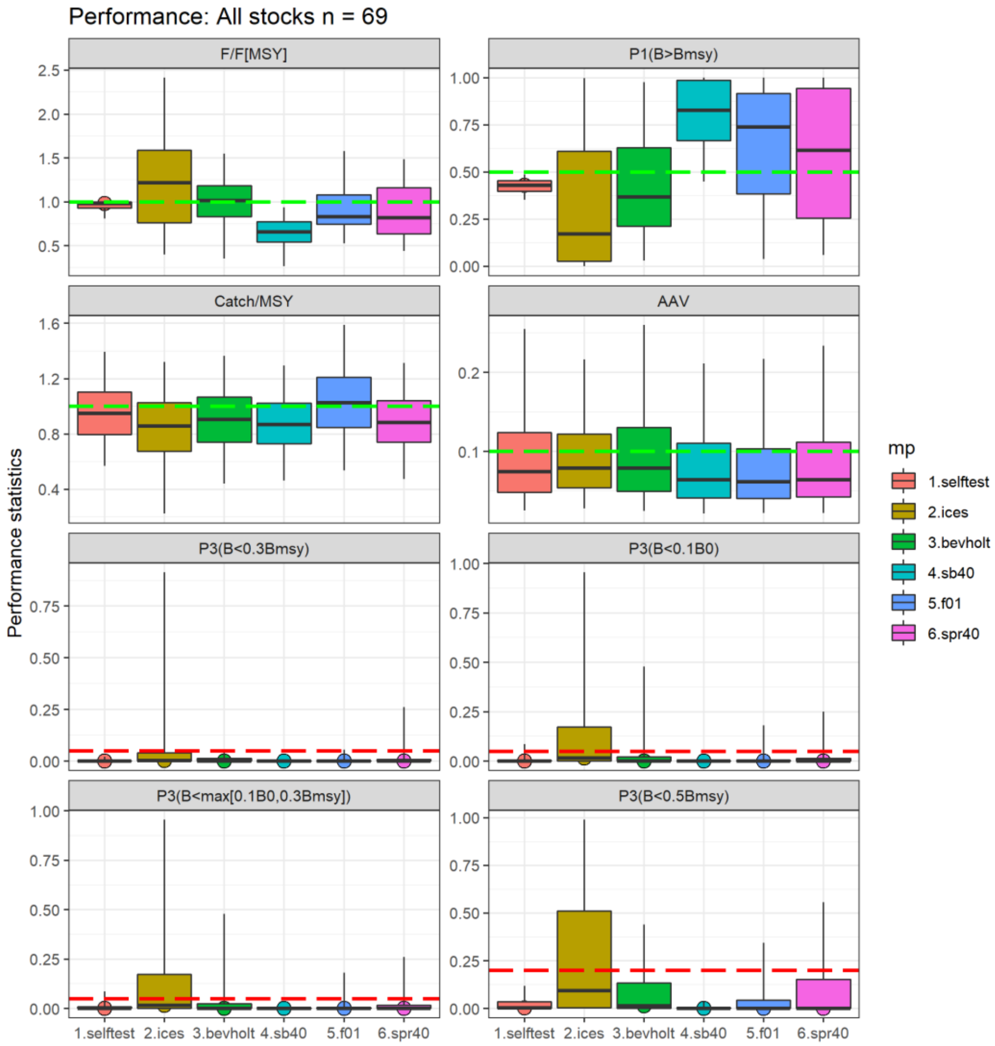


**Figure 4.4. Summary performance statistics of robustness tests for 69 stocks based an evaluation period for 10 years (2071-2080), showing the results for six alternative reference point estimation approaches.**

A comparison of the long-term stock status for the three most common taxonomic orders showed that ICES approach was relatively robust in preventing an over-fished state of clupeiformes, but lead to quite severely overfished state for the majority of stock of the orders gadiformes and pleuronectiformes (Figure 4.5). This pronounced difference appears linked to whether or not $F_{MSY}$ is determined by $F_{MSY} = F_{P.05}$ or to the direct estimate of $F_{MSY} = F_{MMY}$ (Table 4.3). Of the 12 clupeiformes stocks, comprising the foraging species herring, sardines and sprat, all but one herring stock invoked the precautionary $F_{P.05}$. By contrast, gadiformes and pleuronectiformes invoked $F_{MSY} = F_{P.05}$ for only 30% and 25% of stocks, respectively.

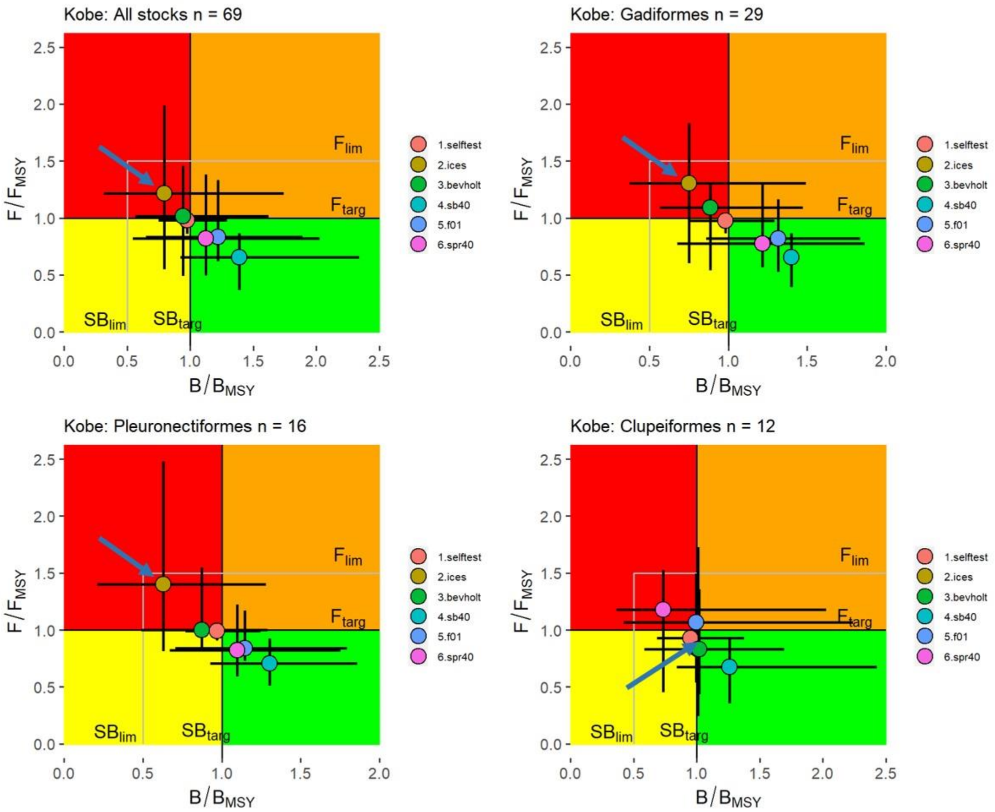


**Figure 4.5. Kobe phase plots showing the relative stock status in terms of $B/B_{MSY}$ and $F/F_{MSY}$ over the evaluation period 10 years (2071-2080) for all 69 stocks and for the three most common taxonomic orders gadiformes, pleuronectiformes and clupeiformes, showing the results for six alternative reference point estimation approaches.**

**Table 4.3. Percentage of stocks where $F_{P.05} < F_{MMY}$ (i.e. $F_{MSY} = F_{P.05}$) for the three most common taxonomic orders.**

| Order | $F_{P.05} < F_{MMY}$ | *N* |
|---|---|---|
| Gadiformes | 30.00% | 29 |
| Pleuronectiformes | 25.00% | 16 |
| Clupeiformes | 91.70% | 12 |

The performance statistics of the 37 stocks (55%) that fall into the group of stocks for which $F_{MMY} = F_{MSY}$ showed further depreciated performance statistics including a very low probability of less 10% to attain $B_{MSY}$, high inter-annual catch variation and a substantially increased risk of breaching the $B_{lim}$ thresholds (Figure 4.6). The group of stocks for which $F_{P.05}$ is invoked for the *ices* $F_{MSY}$ indicated a generally higher risk profile across the other reference point approaches (Figure 4.7). In particular, the *spr40* estimator performed poorer for these stocks. The reason for this is related to dominance of clupeiformes species that tend to undergo larger biomass fluctuations are therefore more risk prone to stochastic depletion (Thorson *et al.*, 2015; Mildenberger *et al.*, 2021). In fact, Mildenberger *et al*. (2021) suggested that for this reason $B_{trigger}$ should be set to $B_{MSY}$ or higher for short-lived species. An unexpected finding of the robustness testing is therefore that the *ices* approach shows improved performance statistics for these stocks, with a higher probability of achieving biomass levels above $B_{MSY}$ and reduced risk of falling below $B_{lim}$ (Figure 4.7). This striking inconsistency is probably related to the combination of more precautionary $F_{MSY}$ values and relatively higher $B_{lim}$ and thus $B_{trigger}$ specifications relative to $B_0$ and $B_{MSY}$ (Figure

3.14) that lead to a more effective reduction in F before biomass levels can decline too close to $B_{lim}$. On the other hand, setting $F_{MSY}$ to typically more risk-prone direct estimates of $F_{MMY}$ in combination with ineffectively low $B_{trigger}$ values (Figure 3.14) can explain the poor robustness of the *ices* approach for stocks that had the lowest risk profiles for all tested alternative approaches. In particular, the use of segmented regressions in the estimation of $F_{MSY}$ results in the stochastic equivalent of the highly risky $F_{max}$ proxy for $F_{MSY}$, because it does not account for the stock–recruitment process for the range of SSB > $B_{lim}$ and may at times be dangerously close to $F_{crash}$ (Mesnil and Rochet, 2010; Punt, 2000; Mace, 2001).

Based on the results from this large-scale simulation testing study, the following list of practical recommendations for improving the robustness of the advice framework is provided:

- Estimating MSY-based reference points directly based on the fitted SRRs (e.g. *ices*, *bevholt*) is associated with high risk over-estimating $F_{MSY}$, poor trade-offs among of SSB < $B_{MSY}$, risk of falling below $B_{lim}$ and yield loss rather than yield gain. In light of uncertainty $F_{MSY}$ proxies should be considered. An alternative is a decrease of $F_{trg}$ relative to the direct estimate $F_{MSY}$ by taking the estimation error into account (Milderberger *et al.*, 2021).
- $F_{MMY}$ estimated involving the segmented regression should not be considered for determining $F_{MSY}$ and thus removed from the ICES toolbox, and, instead replaced by less risk-prone $F_{MSY}$ proxies
- Stochastic forecasts to estimate $F_{MMY}$ must not implement a harvest control rule, because this further increases the risk of overestimating $F_{MSY}$ and somewhat creates a circularity to define a $B_{trigger}$ before the biomass target can be determined.
- $B_{lim}$ should be set to plausible biological biomass levels, at a minimum to $0.1B_0$. If estimated using a hockey-stick, its plausibility should be evaluated against common biological reference points, such $B_{MSY}$ or $B_0$. Alternatively, the "precautionary" conditioned Hockey-Stick formation (Appendix B) produced promising performance statistics for the *f0.1* approach.
- $B_{trigger}$ may be set relative to $B_{lim}$ (e.g. *f0.1*) or $B_{MSY}$ (e.g. *sb40*), but should in any case be guided by $B_{MSY}$ (e.g. *sb40*) to fulfil the minimum condition of, e.g. not being specified below $0.7B_{MSY}$ and equal or higher for short-lived species with high recruitment variability (c.f. Mildenberger *et al.,* 2021).
- There is a clear need for reporting estimates of $B_{MSY}$ or corresponding proxies. When pulling a trigger first, it requires defining a target to aim at, i.e. $B_{MSY}$. Only then is it possible to adjust the trigger if it keeps missing the target. But don't adjust the target.
- The next steps of robustness testing should aim to include generic approaches for estimating $B_{lim}$ and implement the precautionary $F_{P.05}$ rule in all tested advice rules. Different proxies of $F_{MSY}$ should be systematically tested in combination with rules setting $B_{trigger}$ based on a fraction of $B_{MSY}$ or a multiplier of $B_{lim}$ taking into account the biological characteristics of the stocks.

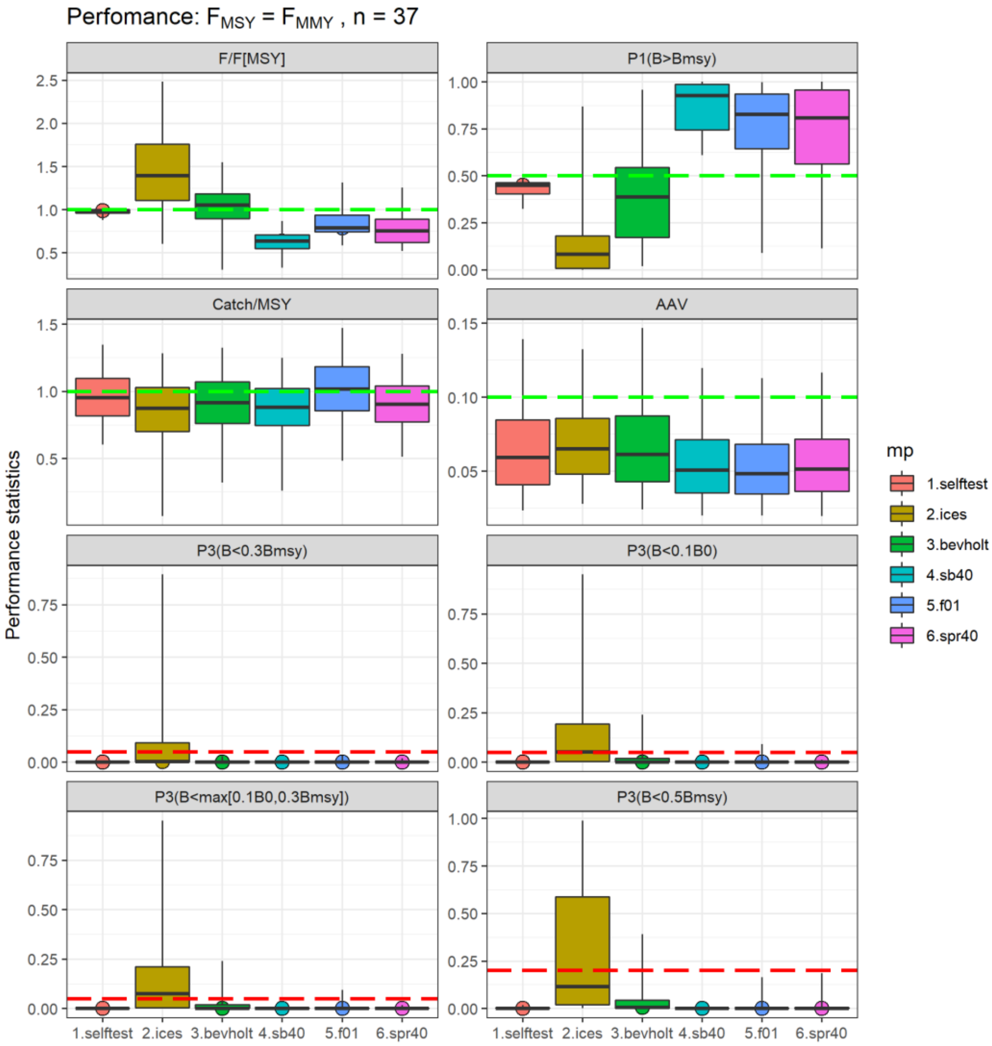


**Figure 4.6. Summary of performance statistics of robustness tests for a group of 37 stocks for which the ICES $F_{MSY}$ was based on the direct estimate of $F_{MMY}$ from the updated EQSIM run based on an evaluation period 10 years (2071-2080), showing the results for six alternative reference point estimation approaches.**

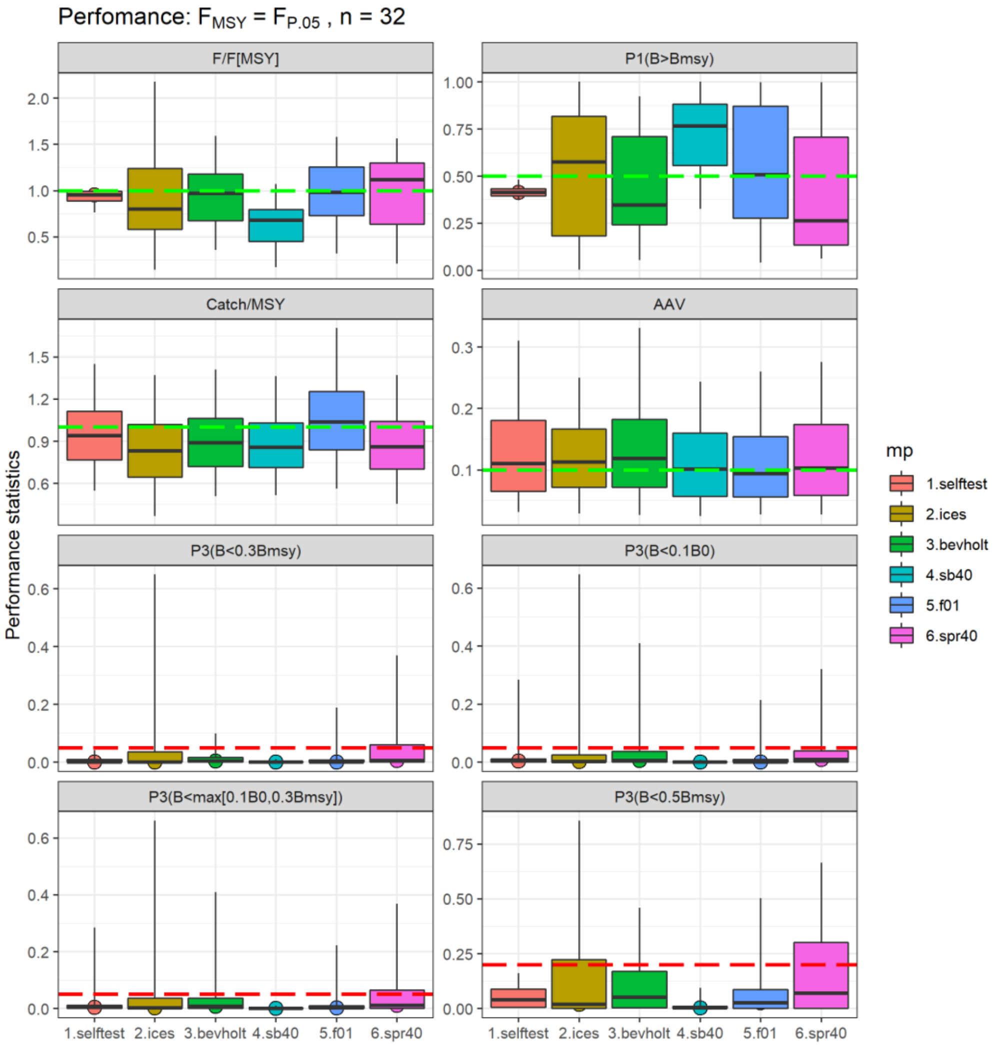


**Figure 4.7. Summary of performance statistics of robustness tests for a group of 32 stocks for which the ICES $F_{MSY}$ was based on the precautionary $F_{P.05}$ from the updated EQSIM run based on an evaluation period 10 years (2071-2080), showing the results for six alternative reference point estimation approaches.**

# 5 Explore alternative methods that can better account for stock dynamics, biological realism and productivity drivers in reference point estimations under climate and environmental uncertainties

## 5.1 Biological realism and exogenous drivers in stock assessment

*by Laurence Kell and Iago Mosqueira*

The ICES Precautionary Approach (PA) and MSY framework requires predicting the results of management actions based on a production function, as used to estimate reference points (Sissenwine and Shepherd, 1987), and forecasts. When providing advice that is robust, in that it can still achieve management objectives despite uncertainty, a key question to ask is: are system dynamics the result of deterministic or stochastic processes (Pennekamp *et al*., 2019)? However, the robustness of advice also depends on process error, due to processes not included in the assessment model. Therefore, permutation entropy, a measure of the complexity of a time series (Bandt and Pompe, 2002) that is negatively correlated with a system's predictability (Garland *et al*., 2018), was evaluated for $SPR_0$, $B_0$ and MSY reference points.

Many stocks were shown to have high entropy as a result of time-varying biological processes, and often this is reflected in empirical data, such as weight-at-age and maturity-at-age information. This inherently translates into non-stationary biological reference of points, such as $F_{MSY}$, $B_{MSY}$ and MSY (Figure 5.1), and may impact both the ability of the model to make accurate forecasts and the ability to manage stocks. Therefore, forecast skill under time-varying conditions and the ability to manage the stock was evaluated using a hindcast approach, by conducting a backtest.

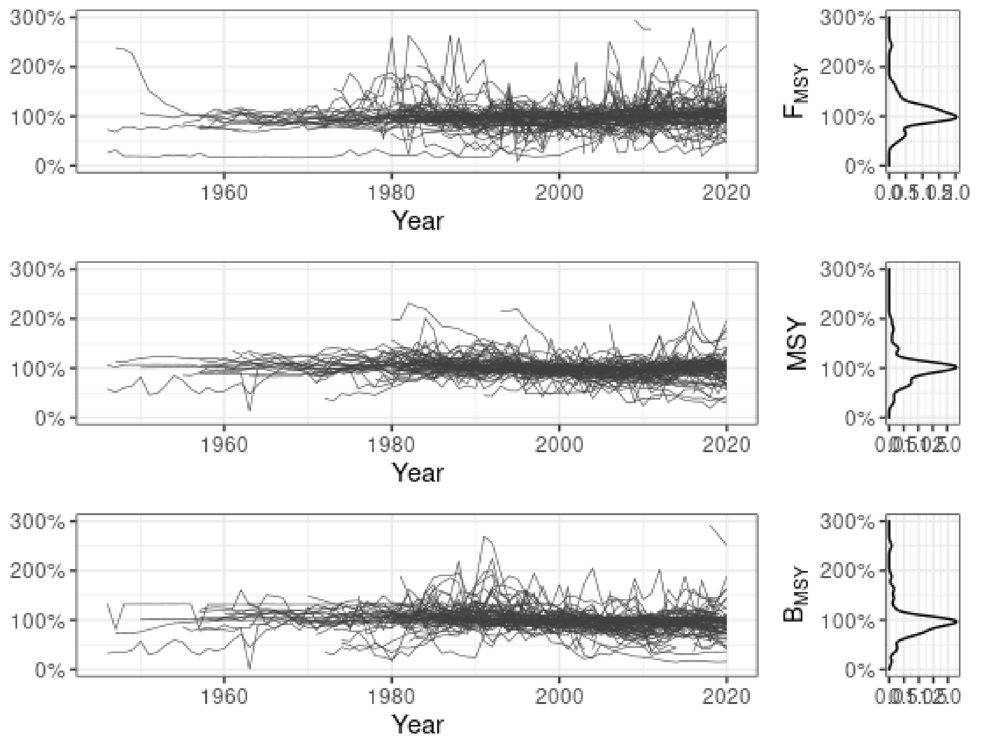


**Figure 5.1. Non-stationarity of biological reference points $F_{MSY}$, $B_{MSY}$ and MSY as a function of annual quantities of weight-at-age, maturity-at-age, natural mortality and selectivity for 69 ICES stocks.**

Forecast skill declined with the distance from the initial conditions (i.e. last years of the assessment). It was also found that some forecasts performed worse than a random walk, had no skill. Since such forecasts are used to set TACs, the reason for this lack of forecast skill should be evaluated, to ensure that management measures are as effective as intended.

A model-based weather forecast should work better than a naive prediction, i.e. saying the weather tomorrow will be the same as today (Kell *et al.*, 2016; Kell *et al.*, 2021; Carvalho *et al.*, 2021). The mean absolute scaled error (MASE) can be used to compare a prediction with a naive prediction, by comparing a prediction of an observation to the corresponding value in a previous time step. A MASE=0.5 means that the forecast is twice as good as a random walk, while MASE>1 means that a random walk is better than the forecast. Hindcasting was carried out by going back 20,19,18, years and forecasting 1,2,3,…,5 years ahead using the reported catch for all the ICES stocks. Prediction skill was evaluated using an "ICES"-type short term forecast procedure with the biology and selectivity represented by an average across 3, 10 and all historical years and recruitment assumed to follow a Beverton-Holt SRR. These scenarios were contrasted by assuming perfect knowledge (2021 assessment estimates) of either biology, selectivity or the recruitment. The results show that the accuracy of the current biology is the determining factor for reliable short-term projections of SSB. Assuming a 3-year average for biological conditions led only to a little depreciation of prediction skill when compared to perfect knowledge. Perfect knowledge of selectivity and recruitment cannot improve the prediction skill if the biology fails to represent the current conditions. Generally, prediction skill started to decline after 3-4 years under all scenarios. Both forecast skill and forecast horizon are expected to be reduced for stocks with short generation times and non-stationary in somatic growth, maturation and survival, i.e. high entropy.

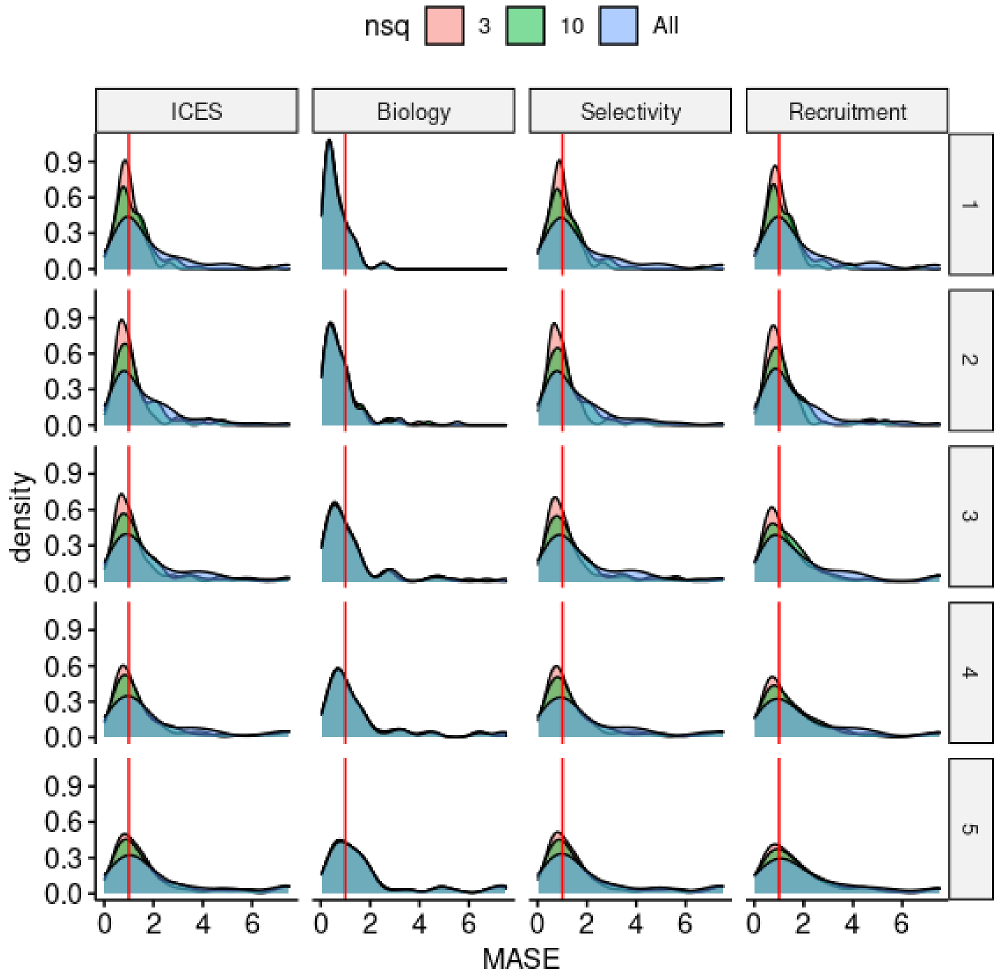


**Figure 5.2. Densities of MASE from the hindcast to evaluate short-term prediction skill over one to five years of forecast (rows). If MASE < 1, the model is that to have prediction skill.**

The backtest, is a form of hindcasting, that allows the impact of a strategy to be evaluated as if it had been used in the past. This makes it easier to understand the potential benefits of adopting alternative strategies. It is also easier to perform than a full MSE, and could be used to filter proposed MPs. The backtest showed that if advice would have been based on 5-years updates on $F_{MMY}$ (corresponding to the maximum median yield from EQSIM) this could have led to increased yield when compared to the actual realized yield over the last 20 years. However, SSB would have still fallen well below $B_{MSY}$ levels (median 0.5-0.6 $B_{MSY}$) that would maximize long-term yield at MSY level. Not updating the 2000 $F_{MMY}$ estimate over the last 20 years would lead to more forgone yield and a further reduction in SSB to levels that are broadly comparable to the current state (Figure 5.3). The performance was highly variable across stocks, which may partially be attributed to recruitment variation. This suggests that $F_{MMY}$ needs to be evaluated on a stock-specific basis, but is generally too high in cases where it is not adjusted by $F_{P.05}$.

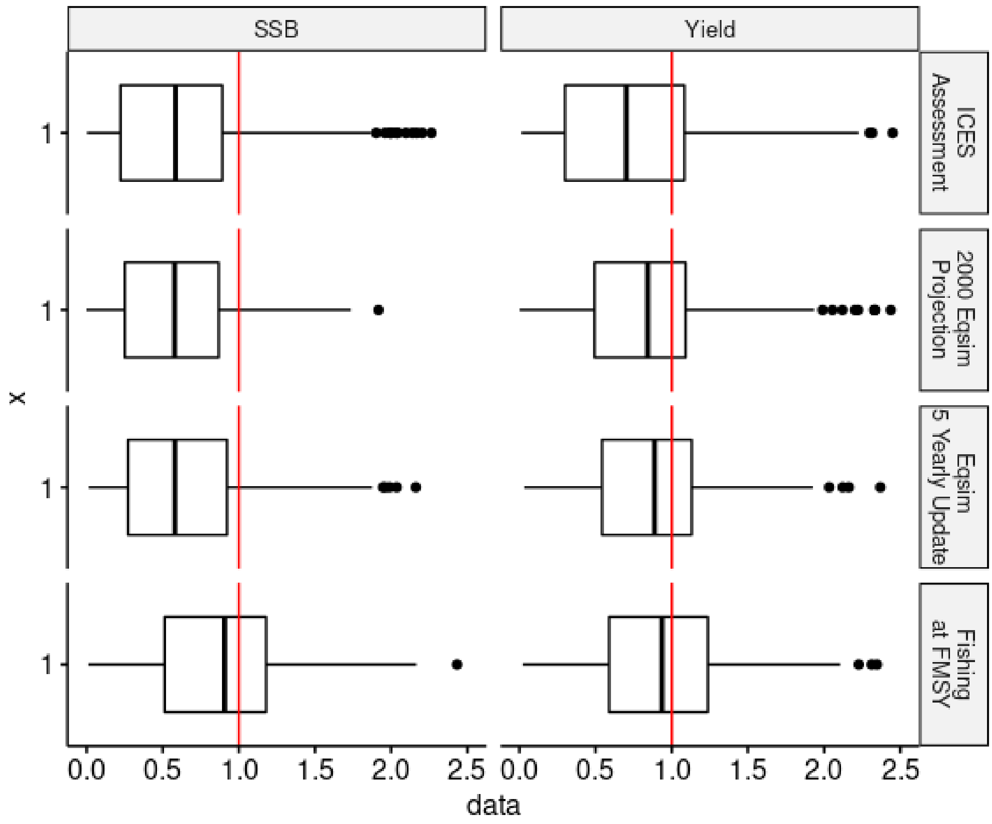


**Figure 5.3. Summary of backtest for all stocks, SSB and yield are relative to time varying MSY reference points for 2016 to 2020.**

The overall conclusions from this analysis are

- Do not use any stock assessment method or procedure used to provide advice (i.e. reference points, forecasts and HCRs) intended to provide either the operational parameters of a management procedure, or the reference points used to evaluate its performance, without simulation testing it first, ideally under the same conditions (e.g. data quality, information content) where it is expected to perform.
- Validate assessments using observations and prediction skill routinely to evaluate the ability of a model-based procedure for providing advice on future catches (Kell et al. 2021). The backtest can be used to filter candidate reference points and HCRs, since if something did not work in the past, why expect it to work in the future? The backtest and hindcasting will also allow the Value-of-Information to be evaluated, i.e. how better data or increase in knowledge under a precautionary approach will improve yield, and so have a direct and measurable economic benefit.
- If multiple models explain the data equally well, then it is important to explore alternative hypotheses and adapt as required. This can be done by using either model ensembles or by conducting Management Strategy Evaluation. When using an ensemble, prediction skill (based on historical observations) can be used to weight models, and forecast skill (based on predicting future state under different management actions) to identify reference points and HCRs used to agree management action. If multiple models can explain the data equally well, MSE should be conducted using multiple Operating Models to ensure that strategies are robust to uncertainty about system processes. To provide robust advice, Management Procedures that use empirical rules and simpler stock assessments, and that make fewer assumptions about the correct processes (e.g. those related to density-dependence), should be evaluated.

- When developing a system there are two important stages: verification and validation. Verification determines if the end product was implemented as proposed, while validation evaluates whether the end product is fit for purpose, i.e. are the PA and MSY objectives met, so that despite uncertainty can long-term sustainable yields be achieved? Once a reference point system has been agreed, tools like hindcasting and backtesting can be used for validation.

## Updating reference points in a changing environment

Biological reference points are a time varying entity, as they are dependent on the productivity of the stock (e.g. weight at age, maturity, natural mortality and selectivity). This implies that, as stated by WKCHANGE, biological reference points should be re-estimated at benchmark assessments, which generally occur on a 5-year cycle. This timescale matches the management system, avoids erratic changes ("whipsaw") in the designation of stock status, and provides some stability in planning horizons for fisheries. The need to update reference points may be more urgent for species with shorter life cycles as there are fewer age classes in the population and hence changes impact the stock more rapidly. On the other hand, short-term variability may occur without trend, in which case reference points can stay the same.

When conducting forecasts, forecast skill declines with the distance from the initial conditions (i.e. last years of the assessment). After 3-5 years, ICES predictions are basically worse than a random walk, thus the system has effectively no prediction power. This again reinforces the concept that biological reference points need to be re-estimated regularly at benchmark assessments and that short-term forecasts should not exceed a three-year time span. The best way to improve prediction skill is to get the biology right and update the biological knowledge with empirical data when those are becoming available.

# 6 Consider appropriate methods of propagating model, estimation and process error uncertainties in the estimation of reference points.

## 6.1 Estimation of reference points in state space models, SAM

*by Christoffer Moesgaard Albertsen*

Reference points are an integral part of fisheries management. Until recently, fish stock assessments using the SAM model had to estimate reference points in post-hoc analyses. However, assumptions made in post-hoc analyses are often not consistent with the original assessment model. Further estimation uncertainty of the estimated reference points is often not quantified. Now, recent developments in the SAM model and R package has allowed estimation of model-consistent reference points with confidence intervals (Albertsen and Trijoulet, 2020; Nielsen *et al.*, 2021). Likewise, the approach to reference point estimation is implemented in the multiStock-assessment package (Albertsen *et al.*, 2018; Albertsen, 2021), which extends the SAM model to multi-stock assessments, and can be implemented for any statistical assessment model where a relevant criterion for the reference point estimation can be calculated.

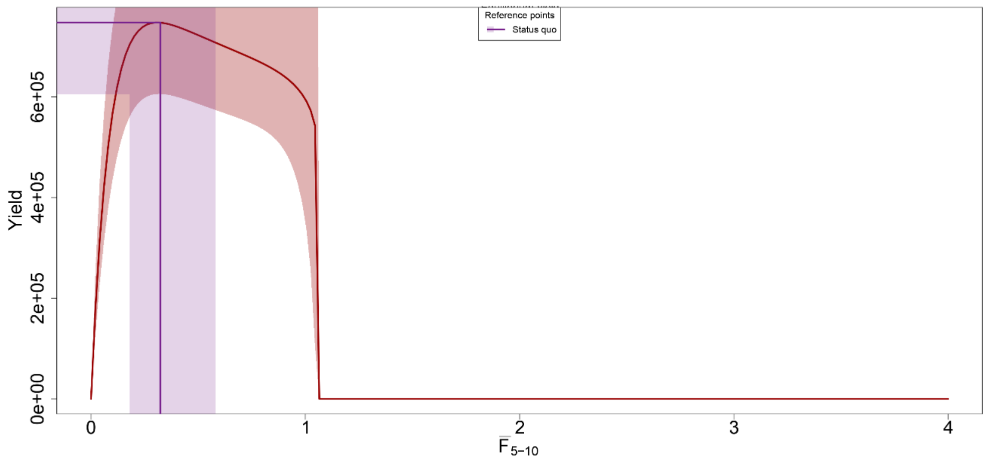


**Figure 6.1. Example of an estimated equilibrium yield curve for Northeast Arctic cod with sigmoidal Beverton-Holt recruitment. The figure is standard output using the plot function on a reference point object from the stockassessment R package.**

The developments were aimed at providing a general, transparent framework for implementing well-defined, model-consistent reference points within the SAM model. As a result, several reference points were implemented, and others can be included with limited implementation effort. As a well-defined, model-consistent reference point, we consider any reference point that can be defined as an optimum of a function calculated based on the fitted assessment model. For example, $F_{max}$ can be defined as the optimum of the yield-per-recruit curve, which can be calculated from the estimated population model. Further, confidence intervals are provided for reference point estimates. The confidence intervals are calculated through a combination of the implicit function theorem and the delta method (Albertsen and Trijoulet, 2020).

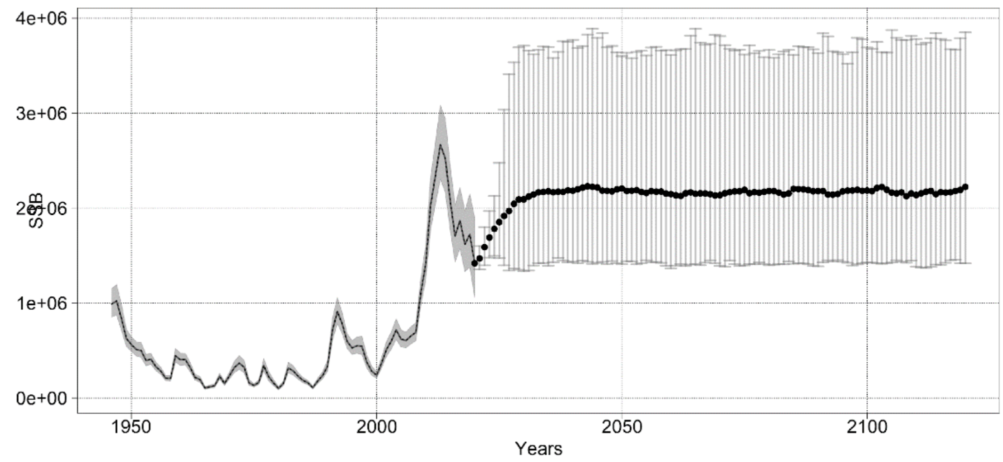


**Figure 6.2. Example of forecasted SSB for northeast Arctic cod by the hcr function in the stockassessment package**

As part of the developments, 11 reference points and 16 recruitment functions - including two-parameter, three-parameter, time series and spline models - are now implemented in the SAM model. The code is structured such that new reference points and recruitment models can be implemented with limited effort. Further, a new forecast module was implemented to provide model consistent forecasts. Forecasts can be made using either the Laplace approximation or simulations. Both can be combined with harvest control rules. The reference points implemented do not currently account for stochasticity in the population model. As such, the reference points should be interpreted in a deterministic context. However, stochastic reference points can be obtained through simulation forecasts. Likewise, probability statements about SSB and related reference points (e.g. MSY $B_{trigger}$ and $F_{P.05}$) can be obtained through simulation-based forecasts. While the reference point estimates are inherently consistent with the assessment model, some assumptions must still be made about the future. In particular, assumptions must be made about the future selectivity and biological input data such as maturity and weight at age. For both selectivity and biological data, an average over a user-defined set of years is used. The new functionality is shown using Northeast Arctic cod as an example. For this example, the assessment model was fitted with a sigmoidal Beverton-Holt stock-recruitment relationship.

**Table 6.1. Reference points implemented. For each, F, equilibrium biomass, yield, recruitment, yield-per-recruit, and biomass-per-recruit is obtained through yield-per-recruit calculations.**

| F reference point | Description |
|---|---|
| **Status Quo** | Last year F |
| **Zero catch** | F=0 |
| **MSY** | Maximum yield |
| **MSY range** | x% of maximum yield (upper and lower) |
| **Max** | Maximum yield per recruit |
| **0.1** | Yield gradient is 10% of gradient at zero |
| **Crash** | Smallest F with $B_e$=0 (from replacement line) |
| **Ext** | Smallest F with $B_e$=0 |
| **x%** | x% of spawners per recruit for F=0 |
| **lim** | $B_{lim}$ for segmented regression, otherwise just before Crash |

## 6.2 Comparison of MSY reference points estimated inside and outside the assessment model

*by Vanessa Trijoulet, Casper W. Berg, David C. M. Miller, Anders Nielsen, Anna Rindorf, and Christoffer M. Albertsen*

Reference points are key to attaining maximum sustainable yield (MSY) and avoiding risk of stock collapse. They can be estimated externally to the assessment model or internally. Both approaches have advantages and disadvantages, but little is known about the effect of choosing one approach over another on reference point estimates. The objective of the study is therefore to compare MSY reference points obtained internally to the assessment model with those obtained by external estimation

Eleven ICES stocks assessed with SAM (Nielsen and Berg, 2014) are retrieved from stockassessment.org and used to provide a basis for the simulations. Three stock-recruitment relationships (SRRs) are fitted to each stock (Ricker, Beverton-Holt and smooth hockey-stick) and the fits are kept for further analysis if converged. This resulted in 27 fits in total. For each fit, MSY reference points were analytically estimated (Albertsen and Trijoulet, 2020) to define the "true" reference points. The fits are then used to simulate 500 replicates, resulting in 13 500 operating models (OMs) in total.

For each OM, an estimation model (EM, SAM) is fitted assuming the 3 SRRs independently and a random walk (RW) on recruitment. Thereafter, MSY reference points are estimated internally and externally using the EMs. Three different estimation approaches are investigated, an internal per-recruit approach (PR) in SAM (Albertsen and Trijoulet, 2020) and two external approaches based on EQSIM (EqS and EqSrw, Simmonds and Millar, 2019). The EqS approach is based on model assumptions that most closely resemble the OMs. The EqSrw approach assumptions follows the ICES guidelines for estimation of reference points which stipulates that the stochasticity in the fish population (e.g. maturity, weights, fishing selectivity) should be taken into account (ICES, 2021). The EqSrw approach relies on the outputs of the stock assessment model assuming RW on recruitment.

MSY estimation is both made with the assumption that the functional form of the SRR is both known and unknown (Figure 6.3). When it is known, the EMs assuming the same SRR than in the OM is used for the PR and EqS approaches. The RW EMs are used for EqSrw. For the external approaches, the same SRR as in the OM is then fitted in EQSIM and used during projections to estimate MSY reference points. When the SRR is unknown, the EM with lowest AIC is used for the PR estimation. For EqS, the EM that has the same SRR assumption as in the OM is used but a weighted average of the 3 SRRs is fitted in EQSIM. For EqSrw the weighted average method is also used but with the outputs of the RW EMs. Different diagnostics are considered to quantify the performance of each estimation approach (Figure 6.3).

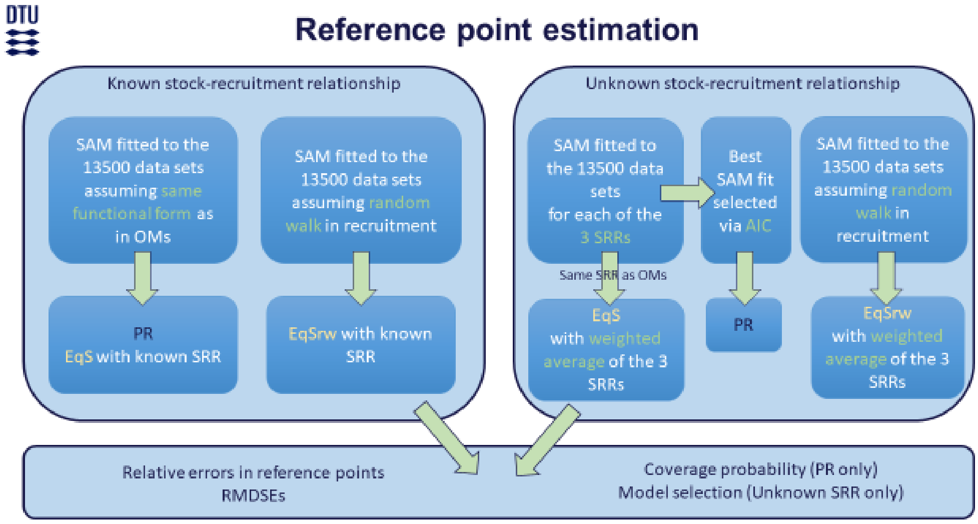


**Figure 6.3. Summary of the methods for reference point estimation and the diagnostics used in the study.**

The study shows that coverage probability of the confidence interval of the reference points is good for the internal approach (not available for the external approaches). Larger variations in bias exist between stocks than between approaches (Figure 6.4), but overall, bias and variance in reference points are lower when estimated internally.

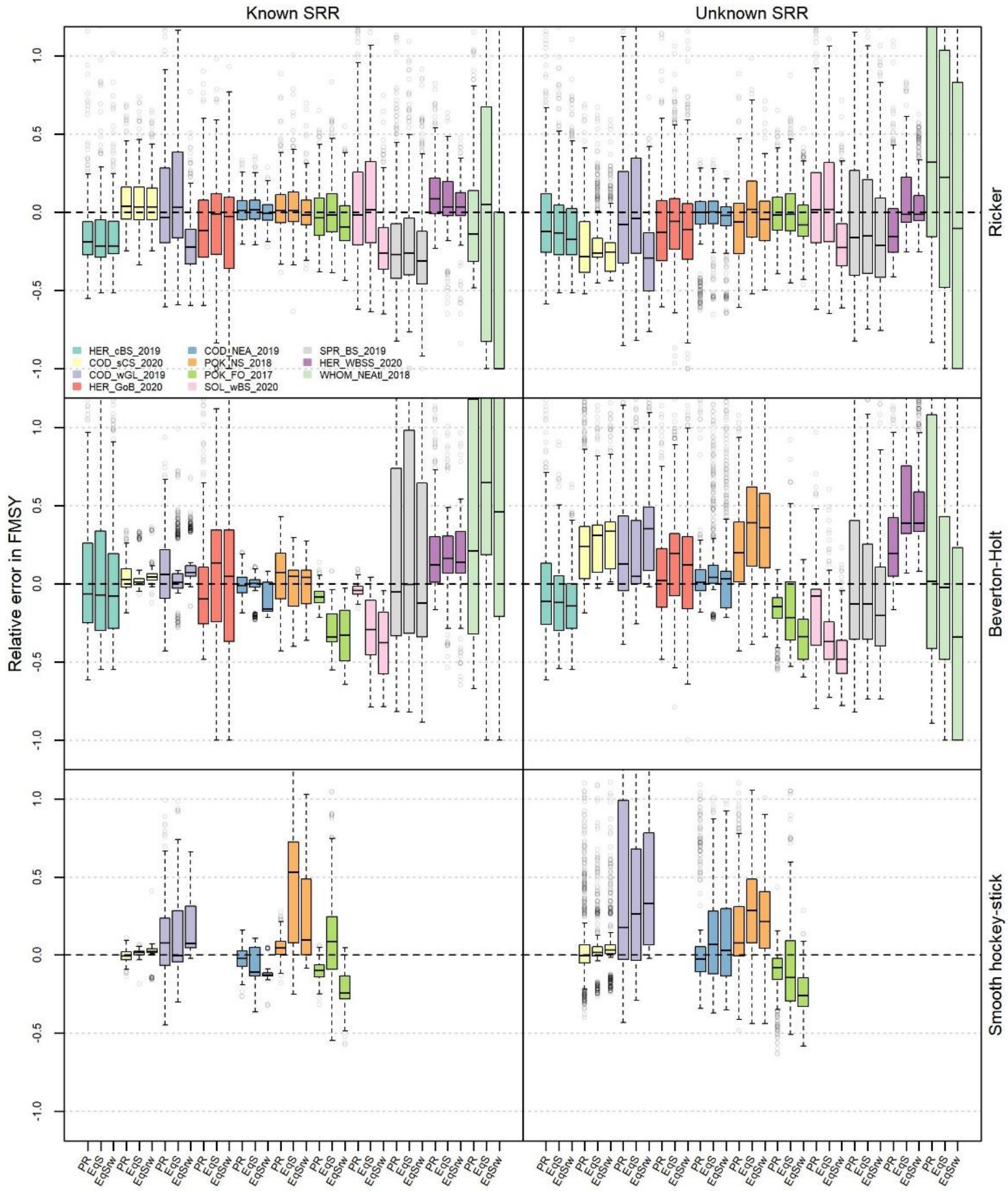


**Figure 6.4. Relative error in $F_{MSY}$ for all stocks and stock-recruitment relationships (SRRs) considered in the study.**

SRR selection is best for the internal approach. When the SRR is wrongly selected, bias can be larger for PR than EqS. This illustrates that using a weighted average of SRRs helps reducing the bias when the SRR is wrongly selected.

The EqSrw is the approach that does the worst for all performance metrics illustrating that stochasticity considerations via re-sampling in the external approach can lead to bias. Further work is therefore needed to investigate the effect of external stochasticity on reference point estimates.

For the ICES stocks assessed using SAM, estimation of the reference points internally can be relevant to use during benchmarks as it is the method that provides the best performance overall and maintains consistency with the assessment model in the estimation of reference points and their confidence intervals.

## 6.3 Reference Points and Projections in Stock Synthesis Model

*by Richard Methot, NOAA Fisheries*

The Stock Synthesis (SS3) model calculates fishery reference points and population projections using a selected reference point of fishing mortality rate (F) level. Incorporating these aspects of stock assessment into an unified package means that all complexities of the historical reconstruction (fleets, areas, discarding, growth estimation, etc.) are also embodied in the reference points and projections. One model run estimates population and fishery parameters based on historical data, then does the reference point calculations, using the spawner-recruitment and fishery parameters, then the projection starting from the end year of the assessment and using the reference point F. This unified approach allows for the variance of estimated population and fishery parameters to influence the variance of reference point and projection quantities. For example, it is routine to get from SS3 the variance of the population projection 10 years into the future and thus to calculate the probability that the population will be above $B_{MSY}$ in each of the projection years, and the covariance of that probability with a parameter such as the estimated steepness of the spawner-recruitment relationship.

The system is designed to implement the management quantities used in the U.S. fishery management system and has proven flexible enough to be useful in other systems. It is strongly tied to the MSY paradigm, but various empirical alternatives are available. SS3 will calculate the equilibrium F that matches a specified Spawner Potential Ratio (SPR), F to match a specified biomass level relative to $B_0$, $F_{0.1}$, and $F_{MSY}$. In all cases, SS3 outputs both per recruit and absolute biomass quantities where the latter accounts for the spawner-recruitment relationship (Figure 6.5).

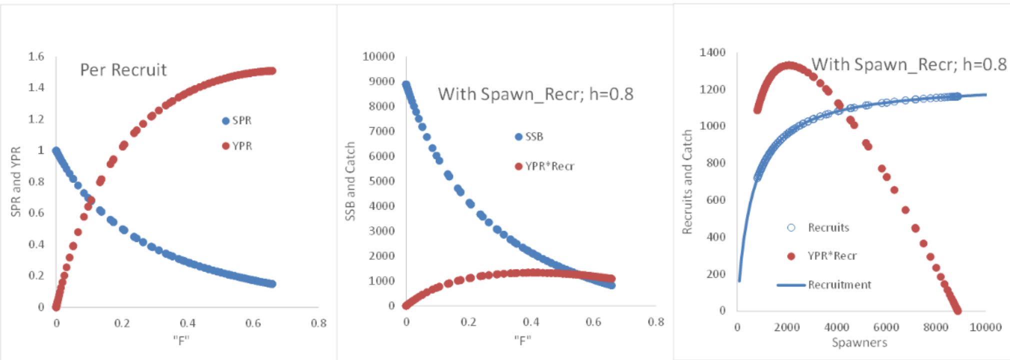


**Figure 6.5. Three panels showing the equilibrium calculations used for reference points. Left panel shows in red dots the Yield Per Recruit (YPR) and Spawning Biomass per Recruit (SSB/R) as a function of F. Each dot is result of equilibrium calculations for that F level. Middle panel takes the spawner-recruitment relationship (here with steepness=0.8) into account and presents spawning biomass and yield (catch) in absolute terms relative to F. The third panel presents yield (catch) and recruitment as a function of spawning biomass.**

The projection (forecast) system in SS3 has great flexibility to deal with a diversity of fleet characteristics. Output is in terms of both catch biomass and catch numbers to enable equivalent quota settings for various fleets. It allows specification of catch ratios (allocations) between fleets or groups of fleets. The flexibility in the projections includes an advanced capability for calculating the expected consequences of holding the fishery to an assessment-informed quota for several years. It does this by conducting the final forecast after two intermediate forecast loops:

1. In the first forecast loop, SS3 uses the F limit level, no recruitment deviations, and no catch allocations or caps. This provides future levels of catch quotas conditioned only on projected stock abundance and fishery selectivity. Future recruitment deviations have

not yet occurred, so it is not correct to forecast expected future catch that assumes knowledge of those future deviations.

2. The second loop uses the target fishing level, termed Acceptable Biological Catch (ABC) in the U.S. This F level is typically specified as a fraction, say 0.75, of the F limit and a biomass-linked inflection and cut-off point are enabled (Figure 6.6). Recruitment deviations and other time-varying parameters are not enabled in this loop because when future catch quotas are set, the future state of the system is not yet known. Catch allocations and caps are then applied and these resultant future fleet-specific catch amounts are stored. In effect, future quotas are being set.
3. In the third loop, the stored catches from loop 2 are brought back into use and the F – catch process is reversed. The F needed to catch that amount, now taking into account recruitment deviations and time-varying parameters is calculated. There also is a provision to treat the actual future catch to have implementation error relative to the quotas that were set when calculating the F needed to catch that amount.

This complex projection process allows SS3 to calculate the expected F, with variance, some years into the future where that F is based on quotas set using the current assessment and a target F policy.

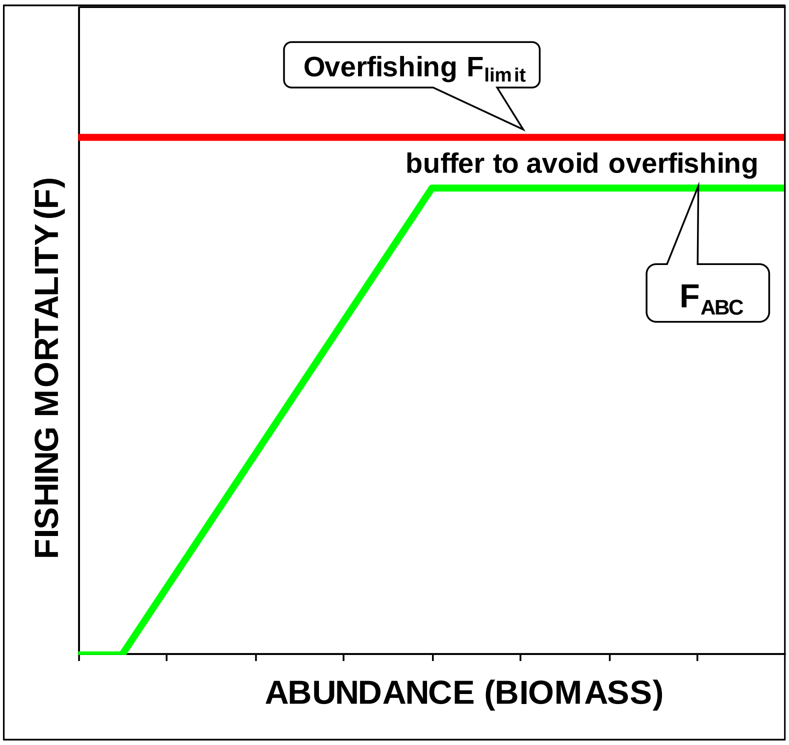


**Figure 6.6. Example of control rules used in the SS3 projections. The red horizontal line shows the fishing mortality limit and the green line with inflection point shows a typical target F level as a function of spawning biomass.**

# 7 Propose candidate methods to address the emerging issues identified under (a) - (d).

It should be stressed that the Reference Points recommended here are based around the ICES advice system, where the Reference Points are revised at each benchmark approximately on a 5-year cycle. Therefore, the Reference Points are valid over that 5-year period and reflect current environmental, stock, and fisheries conditions. They should be taken as values that apply over this limited time period, not as values appropriate for all possible conditions. One consequence of this is that if there are major changes in stock, environment, or fishing it may be necessary to revise the Reference Points, which is usually done in ICES during inter-benchmarks.

### A hierarchical approach to the ICES Reference Point System revision

1. $\mathbf{B_{lim}}$ is integral to the ICES precautionary approach

   - **Type 1**: Consider an empirical Hockey-Stick for deriving $B_{lim}$ only if the data show contrast and a break point is clearly defined
   - **Type 2**: Determine a plausible $B_{lim}/B_0$ ratio based on biological principles and life history of the stock (as for instance 10% to 25% of $B_0$ depending on the type of stocks)
   - **Type 3:** For stocks where the stock development is dominated by occasional good year-classes (i.e. spasmodic recruitment), the lowest observed SSB(s) that gave rise to a good year class can be used as basis for $B_{lim}$
   - Alternative approximations (i.e. current type associated with subjective decisions) should be discouraged
   - Biological plausibility checks (e.g. $B_{lim} > 0.1$ of $B_0$) to ensure there is a sufficient safety margin when setting $B_{lim}$

2. $\mathbf{F_{P.05} \sim f(B_{lim})}$ risk analysis using stochastic projections:

   a) Ideally stochastically projected within the assessment model
   b) If this is not possible, use EQSIM simulator or alternative stochastic projection tools
   c) $F_{P.05}$ should be calculated without $B_{trigger}$

   Note that $B_{lim}$ is strictly an input to the stochastic projections as determined by the benchmark process.

3. $F_{MSY}$ proxy

   a) $\mathbf{F_{MSY}}$ is founded on theory, but challenging to estimate reliably
   b) Proxies for $\mathbf{F_{MSY}}$ should be used consistent with international best practice.
   c) Mainly $\mathbf{F_{spr\%}}$, $F_{B\%}$ (typically 30-50%) and $\mathbf{F_{0.1}}$, where the choice is dependent on the stock's biology.
   d) The $\mathbf{F_{MSY}}$ proxy must not exceed $\mathbf{F_{P.05}}$ without $B_{trigger}$.

4. The **biomass target** ($\mathbf{B_{trg}}$) is the expected mean biomass that corresponds to $F_{MSY}$ as defined in 3 (d)

5. The biomass trigger ($B_{trigger}$)

   a) fraction of $B_{trg}$ (for example 0.7 - 0.9 of $B_{trg}$) or multiplier of $B_{lim}$ (for example 2 x $B_{lim}$) that could be used as $B_{trigger}$ in a new ICES advice rule. Simulations and trade-offs will be presented at WKREF2.

6. $F_{pa}$ and $B_{pa}$ are not needed in the future framework and should be eliminated from the ICES reference points toolbox.

## Justifications:

According to international standards, the limit reference point, $B_{lim}$, and a biological reference point proxy for $F_{MSY}$ (hereafter defined as $F_{brp}$) are the overarching elements of any reference point system, while not exceeding $F_{P.05}$ is a prerequisite for fulfilling the Precautionary Approach. The biomass target ($B_{trg}$) should correspond to $F_{MSY}$= min($F_{P.05}$, $F_{brp}$) and $B_{trigger}$ should provide a sufficient safety buffer to $B_{lim}$, which is more likely and consistently achieved if it is based on a ratio of $B_{trg}$. It was proposed to address these components hierarchically in the following order:

(1) $B_{lim}$

(2) $F_{P.05}$

(3) $F_{brp}$ proxy for $F_{MSY}$

(4) $F_{MSY}$ = min($F_{brp}$,$F_{P.05}$)

(5) $B_{trg}$

(6) $B_{trigger}$

### $B_{lim}$

Ideally $B_{lim}$ should represent the point at which recruitment is reduced. However, computing this directly is generally difficult. Other than directly estimating $B_{lim}$ as the change point of the segmented regression (Type 2 in ICES, 2021a), both the choice of Type and the corresponding estimator have a high degree of subjectivity and can create conflict with the assessment assumptions. For example, $B_{loss}$, which has been used often by ICES to determine $B_{lim}$, is simply a consequence of the history of the exploitation of the stock and has no biological underpinning. Thus, it should be removed from the ICES reference point toolbox (except in cases where there is no realistic SSB-recruitment function, as mentioned below).

Alternative approaches (for example some fraction of $B_0$) also have difficulties: namely in identifying the appropriate fraction, and estimating $B_0$. There is a lot of variability between stocks, which implies that in this case a "least bad" (e.g. $B_{lim}$ always larger than 10% of $B_0$) is going to be the best we can get.

In other words, an absolute $B_{lim}$ should only be specified empirically in cases where there is sufficient contrast in the S-R data to estimate a well-defined break-point, otherwise it is suggested that $B_{lim}$ be specified as a ratio of $B_{MSY}$ or $B_0$ based on biological principles in accordance with international best practice (e.g. 10-25% $B_0$ depending on the stocks characteristic; Sections 7). This is also in line with the principle that $B_{lim}$ should be set at levels of SSB that avoid possible depensation (Allee effect). According to the most recent estimate, the Allee effect for cod occurs when the stock is below 10-25% of $B_0$.

The goal is to arrive at a $B_{lim}$ which marks the point at which recruitment is impaired. In some cases, there will be enough contrast in the data to directly estimate this, in which case $B_{lim}$ should be estimated directly from the data. If there is not enough contrast to estimate $B_{lim}$ directly, then some fraction of $B_0$ can be used, since $B_0$ is estimated from the top of the recruitment function and does not rely on being able to estimate the break point of that recruitment function. We should note that $B_0$ here may not relate to the physical $B_0$, it is purely a numerical output of forecasting the current dynamics in the model. In this case, we need to borrow strength from similar stocks to find the appropriate fraction.

For spasmodic recruiting stocks it is not generally possible to define a realistic SSB-recruitment function, so in those cases $B_{lim}$ should be based on the lowest observed stock size(s) that gave rise to a good year class.

The approach outlined above represents a hierarchy of possible approaches depending on how well the recruitment function can be defined. Estimating $B_{lim}$ directly from the recruitment curve requires good estimates of the breakpoint in the recruitment function. Estimating $B_{lim}$ from some fraction of $B_0$ requires only estimating the overall level in the recruitment function, while the lowest observed stock size with a good year class is a fall back when neither can be estimated (typically due to spasmodic recruitment).

### $F_{P.05}$

The Group agreed that stochastic projections should be used to determine the fishing mortality $F_{P.05}$ that is associated with 5% probability (risk3) that SSB<$B_{lim}$, in accordance with the ICES Precautionary Approach. Code implementing this approach in the FLR platform is now available in the *mse* package as the function computeFp05().

### $F_{MSY}$ proxy

Short-cut MSE-based robustness tests and full MSE applications confirmed that the current way of estimating $F_{MSY}$ proxy might be insufficiently risk adverse. This is in particularly so in cases when $F_{P.05}$ is not adjusting the $F_{MMY}$, which is also supported by back-testing results (Section 5.1). Therefore, the $F_{MSY}$ proxy should be below absolute $F_{MMY}$ (due to the asymmetry in effects of being above and below in a stochastic system and considering robustness to assessment assumption uncertainties). Under-estimating the $F_{MSY}$ proxy when the true value is larger results in small if any loss of catch in the long term and less risk than over-estimating $F_{MSY}$. The current ICES reference point system was the least robust of all tested reference point estimators (Section 4.3). By contrast, $F_{sb40}$ showed the best performance as a $F_{brp}$ proxy for $F_{MSY}$ when robustness was tested within a short-cut MSE simulation framework, followed by $F_{spr40}$ and $F_{0.1}$. However, the results are different depending on the species and life trait history, and ideally it would be best to test different systems to estimate the reference points on the stock-by-stock basis at the benchmark, although it is realised that this might not often possible. The $F_{brp}$ proxy for $F_{MSY}$ should not exceed $F_{P.05}$

### $B_{trg}$

In ICES, biomass targets are not explicitly reported so far, but $B_{MSY}$ is defined as the expected average biomass if the stock is exploited at $F_{MSY}$ with $F_{MSY}$ corresponding to $F_{MSY} = \min(F_{P.05}, F_{MMY})$. Thus, $B_{MSY}$ is implicit in the ICES system and it is consistent with the $B_{trg}$ reference point and it should be reported in the future. It should be noted that ratios (e.g. F or B/$B_{MSY}$) are generally better estimated than absolute values. Therefore, while reporting $B_{MSY}$ is advised, using it as an operational management target might not be advisable for all stock, but should be evaluated on a stock-by-stock basis.

$B_{trigger}$

The biomass trigger, defined in ICES advice rule as MSY $B_{trigger}$ and the limit reference point $B_{lim}$ are often very close and in large majority of cases determined as MSY $B_{trigger} = B_{pa} = 1.4 * B_{lim}$. A different definition of the trigger is needed, including a clear separation from a target reference point. In general, to have a trigger working properly, it should be lower than a biomass target but sufficiently distant from $B_{lim}$ to function as a parachute, e.g. against years of low productivity or regimes.

## Additional suggestions on future procedures to estimate reference points in ICES

- **Transparency**: There is the need for a more transparent, less convoluted process to estimate reference points. The process should also allow for a wider flexibility than perceived currently, with key decisions on how reference points should be estimated to be decided at the benchmark, notwithstanding within the frame of ICES overarching principles.
- **Estimating reference points within the assessment model**: Consensus was that under ideal conditions, the estimation of reference points should occur within an integrated modelling framework at benchmarks. The use of these methods allows for the integration of process uncertainty, parameter uncertainty and observation error into reference points estimation, and produce reference points that are consistent with the assessment model. If this is not possible, stochastic simulators such as EQSim or other analytical tools (e.g. FLBRP) could be used. In either case it should be transparent which life history parameters are based on the full time series, and which on more recent conditions.
- **Integrate historical landings**: In principle, time series should not be trimmed. Historical landings as far back in time as possible, albeit often available, are generally missing from recent assessments. Thus, efforts should be made to present historical landings during the assessment process and/or integrate those, where possible, to give context to the current levels of yield, and thus provide information on scale.
- **Possible use of relative reference points**: Reference points can be also defined as relative instead of as absolute values for assessment models that are able to estimate reference points within the model (e.g. SAM, Gadgets, Stock Synthesis). This implies that reference points can be implicitly updated at each assessment, but F and SSB are also expressed as relative to $F_{MSY}$, $B_{MSY}$, etc. Having relative reference points also allows the use of ensembles and MSE, where different models can be combined (e.g., production models and age structured models). The use of frequent updates to reference points improves management performance on average; however, results may vary according to stocks and thus need to be treated on a per stock basis. Furthermore, frequent updates of reference points allow the inclusion of the most up to date biological information available and avoid the loss of prediction power associated with long time periods between updates. This mirrors the current ICES system with benchmarks and inter-benchmarks, and should be maintained.
- **Integration of density dependence into RPs estimation:** Density-dependent (DD) processes in fish populations may act in different ways. The most commonly considered density-dependent process is recruitment, which is being included in most of the stock recruitment relationships. However, also growth, natural mortality (predation, diseases, etc) and maturity may be affected by density. There has been evidence for DD in empirical biological data for some stocks (e.g., weight-at-age), but a directional trend cannot be generalised to robustly predict the strength of DD given the data based on the current stock levels.

  Although such processes might be expected as the stock approaches high levels (i.e. closer to $B_0$), this may not be realised under fishing aimed to maintain a stock close to

$B_{MSY}$. Because reference points are regularly updated as part of the benchmark cycle, much of the slow changes in stock life history parameters are included in the advice and short-term forecast, and are not expected to substantially change over the lifespan of the reference point. Therefore, for many stocks, as long as benchmarks are conducted every 3-5 years and the current dynamics are well represented in the assessment model, explicit modelling of the density dependence is not required as a routine component of the default advice rule. However, if there is evidence that DD is an important driver at current/expected stock sizes, then this should be included through a specific analysis and the generic guidelines may not be applicable in this case.

Such cases are most likely to arise in strongly cannibalistic species, in stocks well above $B_{MSY}$ and approaching carrying capacity, or in stocks which have rapid changes in biomass associated with impacts over the life span of the reference points. Rather than attempt to give generic guidelines for these case specific examples, the recommendation is for DD of these stocks to be incorporated into the assessment model, and potentially be handled through specifically designed HCR evaluation/Management Strategy Evaluation.

## Additional recommendations

- WKREF2 should specify how to set the $F_{MSY}$ proxy to be used in the new ICES harvest control rule. It should be species-dependent, based on its biology, resilience and ecology, following international best practice. This can be done through simulations to be presented at WKREF2, or alternatively following international best practice, such as, for example, adopting with or without modifications the NZ system.
- WKREF2 should define which fraction of $B_{tgt}$ (for example 0.7 - 0.9 of $B_{tgt}$) or multiplier of $B_{lim}$ (for example 2 x $B_{lim}$) should be used as $B_{trigger}$ in the new system. This should be done through simulations to be presented at WKREF2.
- WKREF2 should specify the level of probability which is associated with being above (or below) $B_{trg}$ in the new system.
- WKREF2 should define how to estimate reference points based on an MSE framework, when an MSE is conducted for a particular stock. It is important to note that MSE should be used to define the operational arguments of a Management Procedure (MP), while the assessment should calculate the biological reference points, and thus in principle they should not be in conflict. Probability-based metrics, such as $F_{P.05}$, can be derived from the MSE framework.
- Within TAF, it is recommended to generate FLStock objects of all final assessments with the attributes that were provided by WKREF1. In cases where probabilistic statements are already included in the advice, the FLR object should also contain the associated uncertainty. This could also be used to test the impact of assessment uncertainty on e.g. estimates of $F_{P.05}$, and potentially to include it in its estimation.
- WKREBUILD2 should investigate how the outcomes of WKREF1 and 2 affect the evaluation and guidelines of rebuilding plans in ICES, including rebuilding triggers and targets.

# References

Albertsen, C. M., Nielsen, A., and Thygesen, U. H. (2018) Connecting single-stock assessment models through correlated survival. ICES Journal of Marine Science. 75(1). 235-244 doi: 10.1093/icesjms/fsx114

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. doi: https://doi.org/10.1016/j.fishres.2020.105618.

Albertsen, C. M. (2021). multiStockassessment: Fitting Multiple State-Space Assessment Models. R package version 0.3.0. https://github.com/calbertsen/multi_SAM

Bandt, Christoph, and Bernd Pompe. 2002. "Permutation Entropy: A Natural Complexity Measure for Time Series." Physical Review Letters 88 (17). APS: 174102.

Barrowman, N. J., & Myers, recruitment. A. 2000. Still more spawner-recruitment curves: the hockey stick and its generalizations. Canadian Journal of Fisheries and Aquatic Sciences, 57(4), 666-676.

Beverton, R. 1998. Fish, Fact and Fantasy: a Long View. Reviews in Fish Biology and Fisheries, 8: 229–249. https://doi.org/10.1023/A:1008888411100.

Buckland, S.T., K. P. Burnham and N. H. Augustin, 1997. Model Selection: An Integral Part of Inference. Biometrics, Vol. 53, No. 2 (Jun., 1997), pp. 603-618 (16 pages)

Carvalho, F., Winker, H., Courtney, D., Kapur, M., Kell, L., Cardinale, M., Schirripag, M., Kitakado, T., Yemane, D., Piner, K.R., Maunder, M.N., Taylor, I., Wetzel, C.R., Doering, K., Johnsonm, K.F., Methot, R.D., 2021. A Cookbook for Using Model Diagnostics in Integrated Stock Assessments. Fish. Res. 240, 105959.

van Deurs, M., Brooks, M. E., Lindegren, M., Henriksen, O., & Rindorf, A. 2021. Biomass limit reference points are sensitive to estimation method, time-series length and stock development. Fish and Fisheries, 22(1), 18-30.

DFO. 2009. A fishery decision-making framework incorporating the precautionary approach. https://www.dfo-mpo.gc.ca/reports-rapports/regs/sff-cpd/precaution-eng.htm.

Fischer, S.H., De Oliveira, J.A.A., Mumford, J.D., Kell, L.T., 2021. Using a genetic algorithm to optimize a data-limited catch rule. ICES J. Mar. Sci. 78, 1311–1323.

Garland, Joshua, Tyler R Jones, Michael Neuder, Valerie Morris, James WC White, and Elizabeth Bradley. 2018. "Anomaly Detection in Paleoclimate Records Using Permutation Entropy." Entropy 20 (12). Multidisciplinary Digital Publishing Institute: 931.

ICES. 2014. Report of the Workshop to consider reference points for all stocks (WKMSYREF2), 8-10 January 2014, ICES Headquarters, Copenhagen, Denmark. ICESCM 2014/ACOM:47. 91 pp.

ICES. 2015. Report of the Joint ICES-MYFISH Workshop to consider the basis for FMSY ranges for all stocks (WKMSYREF3), 17–21 November 2014, Charlottenlund, Denmark. ICES CM 2014/ACOM:64. 156 pp.

ICES. 2017. Report of the Workshop to consider FMSY ranges for stocks in ICES categories 1 and 2 in Western Waters (WKMSYREF4), 13–16 October 2015, Brest, France. ICES CM 2015/ACOM:58. 187 pp.

ICES. 2019. Workshop on North Sea stocks Management Strategy Evaluation (WKNSMSE). ICES Scientific Reports. 1:12. 378 pp. http://doi.org/10.17895/ices.pub.5090.

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

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

ICES. 2020c. Workshop on guidelines and methods for the evaluation of rebuilding plans (WKREBUILD). ICES Scientific Reports. 2:55. 79 pp. http://doi.org/10.17895/ices.pub.6085. https://doi.org/10.17895/ices.pub.6085

ICES. 2021a. ICES fisheries management reference points for category 1 and 2 stocks; Technical Guidelines. In Report of the ICES Advisory Committee, 2021. ICES Advice 2021, Section 16.4.3.1. https://doi.org/10.17895/ices.advice.7891.

ICES. 2021b. Workshop of Fisheries Management Reference Points in a Changing Environment (WKRP-Change, outputs from 2020 meeting). ICES Scientific Reports. 3:6. 39 pp. https://doi.org/10.17895/ices.pub.7660

ICES. 2021c. The Workshop for the evaluation of the Iberian sardine HCR (WKSARHCR). ICES Scientific Reports. 3:49. 115 pp. https://doi.org/10.17895/ices.pub.7926.

ICES. 2021d. Benchmark Workshop on the development of MSY advice for category 3 stocks using Sur-plus Production Model in Continuous Time; SPiCT (WKMSYSPiCT). ICES Scientific Reports. 3:20. 317 pp. https://doi.org/10.17895/ices.pub.7919

Hilborn, R. 2010. Pretty good yield and exploited fishes. Marine Policy, 34: 193–196.

Hilborn, R., Hively, D. J., Jensen, O. P., and Branch, T. A. 2014. The dynamics of fish populations at low abundance and prospects for rebuilding and recovery. ICES Journal of Marine Science, 71: 2141–2151. https://doi.org/10.1093/icesjms/fsu035.

Hilborn, R. 2020. Measuring fisheries management performance. ICES Journal of Marine Science, 77: 2432–2438. https://doi.org/10.1093/icesjms/fsaa119.

Hordyk, A. R., Huynh, Q. C., and Carruthers, T. R. 2019. Misspecification in stock assessments: Common uncertainties and asymmetric risks. Fish and Fisheries, 20: 888–902.

Johnson, K.F., Councill, E., Thorson, J.T., Brooks, E., Methot, R.D., Punt, A.E., 2016. Can autocorrelated recruitment be estimated using integrated assessment models and how does it affect population forecasts? Fish. Res. 183, 222–232.

Kell, L.T., Mosqueira, I., Grosjean, P., Fromentin, J-M., Garcia, D., Hillary, R., Jardim, E., Mardle, S., Pastoors, M.A., Poos, J.J., Scott, F., Scott, R.D. 2007. FLR: an open-source framework for the evaluation and development of management strategies. ICES Journal of Marine Science, 64: 640-646.

Kell, L.T., Nash, R.D.M., Dickey-Collas, M., Mosqueira, I., Szuwalski, C., 2016. Is spawning stock biomass a robust proxy for reproductive potential? Fish Fish. 17, 596–616. https://doi.org/10.1111/faf.12131

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

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.

Mildenberger, T. K., Berg, C. W., Kokkalis, A., Hordyk, A. R., Wetzel, C., Jacobsen, N. S., Punt, A. E., et al. 2021. Implementing the precautionary approach into fisheries management: Biomass reference points and uncertainty buffers. Fish and Fisheries: 1–20.

Miller, T. J., and Brooks, E. N. 2021. Steepness is a slippery slope. Fish and Fisheries, 22: 634–645.

Ministry of Fisheries. 2008. Harvest Strategy Standard for New Zealand Fisheries. New Zealand Government. https://fs.fish.govt.nz/Page.aspx?pk=113&dk=16543

Ministry of Fisheries. 2011. Operational Guidelines for New Zealand's Harvest Strategy Standard. New Zealand Government. https://fs.fish.govt.nz/Doc/22847/Operational_Guidelines_for_HSS_rev_1_Jun_2011.pdf.ashx

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.

Nielsen, A., Berg, C. W., Albertsen, C. M., Kristensen, M., Brooks, M., Trijoulet, V. and Brevik, O. N. (2021). stockassessment: State-Space Assessment Model. R package version 0.11.0. https://github.com/fishfollower/SAM

Pennekamp, Frank, Alison C Iles, Joshua Garland, Georgina Brennan, Ulrich Brose, Ursula Gaedke, Ute Jacob, et al. 2019. "The Intrinsic Predictability of Ecological Time Series and Its Potential to Guide Forecasting." Ecological Monographs 89 (2). Wiley Online Library: e01359.

Perälä T, Kuparinen A. 2017. Detection of Allee effects in marine fishes: analytical biases generated by data availability and model selection. doi/full/10.1098/rspb.2017.1284

Perälä T, Kuparinen A., Jeffrey Huthchings 2021. Allee Effects and the Allee-effect Zone in Atlantic Cod. Biology Letter, in press

Punt, A. E., Butterworth, D. S., de Moor, C. L., De Oliveira, J. A. A., and Haddon, M. 2015. Management strategy evaluation: best practices. Fish and Fisheries, 17: 303–334. http://doi.wiley.com/10.1111/faf.12104.

Ralston, S., Punt, A. E., Hamel, O. S., Devore, J. D., & Conser, R. J. (2011). A meta-analytic approach to quantifying scientific uncertainty in stock assessments. Fishery Bulletin, 109, 217–231.

Simmonds, J., Hjorleifsson, E., and Millar, C. 2019. msy: Estimation of Equilibrium Reference Points for Fisheries. URL http://github.com/ices-tools-prod/msy. R package version 0.1.19.

Sissenwine, MP, and JG Shepherd. 1987. "An Alternative Perspective on Recruitment Overfishing and Biological Reference Points." Can. J. Fish. Aquat. Sci. 44 (4). NRC Research Press: 913–18.

Thorson, J. T., Jensen, O. P., and Hilborn, R. 2015. Probability of stochastic depletion: an easily interpreted diagnostic for assessment modelling and fisheries management. ICES Journal of Marine Science, 72: 428–435.

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. John Wiley & Sons, Ltd. https://doi.org/10.1111/faf.12427.

# Annex 1: FLSRTMB: Fitting conditioned Stock Recruitment Relationships (SRR) in FLR

***Beverton-Holt SSR conditioning with prior information for steepness***

The stock-recruitment relationship (SRR) was assumed to follow a Beverton and Holt model (BH-SRR) of the form

$$R_y = \frac{aSB_{y-a_{min}}}{b + SBB_{y-a_{min}}} e^{\epsilon_y - 0.5\sigma_r^2}$$

where $R_y$ is the number of recruits in year $y$, $SSB_{y-a_r}$ is the spawning biomass in year $y$ minus minimum age $a_{min}$ defined for the stock (typically age-0 or age-1). The recruitment deviation $\epsilon_t$ is assumed to be associated with a a first-order autocorrelation (AR1) process (Johnson et al. 2016; Simmonds et al. 2019), such that

$$\epsilon_y = \rho\epsilon_{y-1} + \delta_y\sqrt{1-\rho^2}$$

where $\rho$ is the AR1 coeffient and $\delta_y \sim N(0, \sigma_r)$ determines variation in recruitment as a function of the recruitment standard deviation $sigma_r$.

The BH-SRR was fitted the recruitment $R$ and $SSB$ from FLStock objects using the FLR library FLSRTMB (Winker and Mosquiera; https://github.com/flr/FLSRTMB), which enables straightforward integration of available prior information on the steepness $s$ of the SSR from a recent meta-analysis (Thorson 2020).

For this purpose, the Beverton-Holt equation in FLSRTMB is re-parameterised as function of steepness $s$ and annual unfished spawning biomass per-recruit $SPR_0$ (Mace and Doonan, 1988),

$$R_y = \frac{4sSB_{y-a_{min}}R0}{R_0 SPR_{0_y}(1-s) + SBB_{y-a_{min}}(5s-1)}$$

where steepness $s$ is defined as the ratio of recruitment when $SSB$ equals 20% of the unfished $SSB_0$ to the virgin recruitment $R_0$ at $SSB0$. A notable difference to the conventional parameterization is that $SPR_{0_y}$ is treated as non-stationary, being function of annual quantities of $W_{a,y}$, $Mat_{a,y}$ and $M_{a,y}$. By way of using time-varying $SPR_{0,y}$, also takes into consideration the recent criticism by Miller and Brooks (2021) that specifying a set biological parameters to define a single time-invariant $SPR_0$ can be highly sensitive to reference estimation when using steepness values from meta-analysis.

The prior distribution for $s$ is generated from truncated logit distributions ($TrunkLogit$) of the form

$$s = 0.2001 + 0.7999/\left(1 + exp\left(-s_{logit}\right)\right)$$

$$s \sim TrunkLogit\left(s_{logit}, \sigma_{logit}\right)$$

where $s_{logit}$ and $sigma_{logit}$ correspond to the input of species-specific predictions for the distribution of $s$ from the hierarchical taxonomic FishLife model (Thorson, 2020, https://github.com/James-Thorson-NOAA/FishLife), summarized in Table A1. The default prior is assuming an approximately uniform prior between 0.3 – 0.9, with a decreasing density (soft bounds) to the limits 0.2 and 1.0 (Figure. A1)

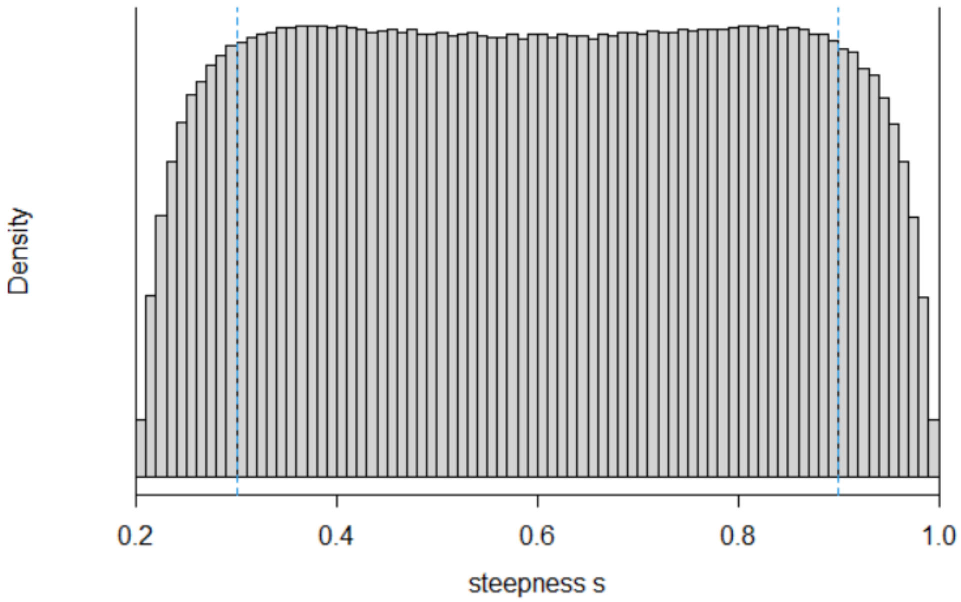


**Figure A.1.1 Graphical illustration of default prior for estimating steepness s, with a mean of 0.6 and logit.sd = 1.5**

The FLSRTMB estimates of $R0$ and $s$ are then converted into the parameters $a$ and $b$ of the Beverton-Holt formulation in FLR, such that

$a = \frac{4sR_0SPR_0}{5sSPR_0-1}$ and $b = \frac{R_0SPR_0(1-s)}{5s-1}$

where the reference for $SPR_0$ to predict $a$ and $b$ was taken the average $SPR_{0_y}$ across all years in the case of the OM.

### *A conditioned, continuous hockey-stick SSR*

A new conditional Hockey-Stick formulation was developed and implemented in 'FLSRTMB`. The new Hockey-Stick is based on a continuous, quadratic hockey-stick (c.f. Barrowman and Myers), which is re-parameterised as a function of $SPR_{0_y}$ and a "re-purposed" steepness parameter $s^*$ given by

$$R_y = \frac{s^*}{2P_{lim}SPR0_y}\left(SSB_y + P_{lim}R_0SPR_{0_y}/s^* - \sqrt{\left(SSB_y - P_{lim}R_0SPR_{0,y}/s^*\right)^2}\right)$$

In addition, the parameter $P_{lim}$ is introduced, which then determines the lower of the ratio $B_{lim}/SSB_{0_y}$, where $Blim$ corresponds to break point $b$ of the segmented regression and $SSB_{0_y}$ is allowed to be treated as non-stationary being a function of $SSB_{0_y} = R_0SPR_{0_y}$.

The break point $b$ ($B_{lim}$) and slope $a$ are given by

$b = P_{lim} * R_0 SPR_{0,y}/s$ and $a = R_0/b$

In the chosen setting for FLSRTMB, the parameter $s^*$ was bounded by a mostly uniform distribution between $0.2 > s^* \leq 1$, with soft bounds towards the limits (Fig. A1). This invokes for the setting $P_{lim} = 0.1$, that $B_{lim}$ is not defined for $0.1B_0 < B_{lim} < 0.5B_0$, while for $P_{lim} = 0.05$, $B_{lim}$ is not defined for $0.05B_0 < B_{lim} < 0.25B_0$.

**Table A.1.1. List Species arranged by taxonomic order with FishLife (Thorson 2020) predictions for the recruitment standard deviation ($\sigma_R$), the auto-correlation coefficient (ρ), steepness (s) and the associated standard error ($\sigma_S$) on logit scale.**

| Species | Order | $\sigma_R$ | ρ | s | $\sigma_s$ |
|---|---|---|---|---|---|
| *Argentina silus* | Argentiniformes | 0.69 | 0.38 | 0.52 | 1.14 |
| *Clupea harengus* | Clupeiformes | 0.67 | 0.32 | 0.58 | 0.26 |
| *Sardina pilchardus* | Clupeiformes | 0.49 | 0.50 | 0.77 | 0.60 |
| *Sprattus sprattus* | Clupeiformes | 0.70 | 0.31 | 0.80 | 0.67 |
| *Brosme brosme* | Gadiformes | 0.42 | 0.56 | 0.57 | 1.30 |
| *Gadus morhua* | Gadiformes | 0.53 | 0.39 | 0.79 | 0.22 |
| *Melanogrammus aeglefinus* | Gadiformes | 0.80 | 0.24 | 0.66 | 0.34 |
| *Merlangius merlangus* | Gadiformes | 0.64 | 0.31 | 0.71 | 0.43 |
| *Merluccius merluccius* | Gadiformes | 0.23 | 0.67 | 0.56 | 1.20 |
| *Micromesistius poutassou* | Gadiformes | 0.60 | 0.34 | 0.55 | 0.73 |
| *Molva molva* | Gadiformes | 0.38 | 0.56 | 0.53 | 1.33 |
| *Pollachius virens* | Gadiformes | 0.46 | 0.57 | 0.79 | 0.40 |
| *Pandalus borealis* | Crustacian | 0.28 | 0.27 | 0.84 | 0.30 |
| *Lophius piscatorius* | Lophiiformes | 0.30 | 0.88 | 0.92 | 1.28 |
| *Dicentrarchus labrax* | Perciformes | 0.34 | 0.75 | 0.90 | 1.93 |
| *Trachurus trachurus* | Perciformes | 0.53 | 0.47 | 0.75 | 0.87 |
| *Glyptocephalus cynoglossus* | Pleuronectiformes | 0.53 | 0.47 | 0.63 | 1.04 |
| *Lepidorhombus boscii* | Pleuronectiformes | 0.37 | 0.68 | 0.87 | 1.23 |
| *Lepidorhombus whiffiagonis* | Pleuronectiformes | 0.38 | 0.66 | 0.84 | 1.29 |
| *Pleuronectes platessa* | Pleuronectiformes | 0.48 | 0.58 | 0.82 | 0.40 |
| *Scophthalmus maximus* | Pleuronectiformes | 0.60 | 0.48 | 0.86 | 1.15 |
| *Solea solea* | Pleuronectiformes | 0.54 | 0.34 | 0.61 | 0.42 |
| *Scomber scombrus* | Scombriformes | 0.78 | 0.28 | 0.64 | 0.58 |
| *Sebastes norvegicus* | Scorpaeniformes | 0.56 | 0.61 | 0.58 | 0.96 |

# Annex 2: List of participants

| Member | Dept/Institute | Email |
|---|---|---|
| Alessandro Orio | Swedish University of Agricultural Sciences | alessandro.orio@slu.se |
| Andreia Silva | Portuguese Institute for the Sea and the Atmosphere | avsilva@ipma.pt |
| Andrés Uriarte | AZTI-Tecnalia | auriarte@azti.es |
| Cecilie Kvamme | Institute of Marine Research | cecilie.kvamme@hi.no |
| Ching Villanueva | Centre de Bretagne | Ching.Villanueva@ifremer.fr |
| Christoph Konrad | Joint Research Centre | christoph.konrad@ec.europa.eu |
| Christoffer Moesgaard Albertsen | DTU Aqua, National Institute of Aquatic Resources | cmoe@aqua.dtu.dk |
| Cóilín Minto | Galway-Mayo Institute of Technology | Coilin.Minto@gmit.ie |
| Colm Lordan | International Council for the Exploration of the Sea | colm.lordan@ices.dk |
| Daniel Duplisea | Institut Maurice-Lamontagne | Daniel.duplisea@dfo-mpo.gc.ca |
| Daniel Howell | Institute of Marine Research | daniel.howell@hi.no |
| David Gilljam | SLU Department of Aquatic Resources Institute of Marine Research | david.gilljam@slu.se |
| David Miller | International Council for the Exploration of the Sea | david.miller@ices.dk |
| David Reid | Marine Institute | david.reid@marine.ie |
| Debra Lambert | NOAA Fisheries | deb.lambert@noaa.gov |
| Dorleta Garcia | AZTI-Tecnalia | dgarcia@azti.es |
| Francesco Masnadi | University of Bologna | francesco.masnadi@studio.unibo.it |
| Gjert Endre Dingsør | Norwegian Fishermen's Association | gjert@fiskebat.no |

| Member | Dept/Institute | Email |
|---|---|---|
| Henning Winker (Chair) | Joint Research Centre | henning.WINKER@ec.europa.eu |
| Henrik Sparholt | University of Copenhagen | henrik.sparholt@gmail.com |
| Iago Mosqueira | Wageningen University & Research | iago.mosqueira@wur.nl |
| Jan Horbowy | National Marine Fisheries Research Institute | horbowy@mir.gdynia.pl |
| Jean-Baptiste Lecomte | Centre Atlantique | jean.baptiste.lecomte@ifremer.fr |
| Johan Lövgren | SLU Department of Aquatic Resources-SLU Aqua | johan.lovgren@slu.se |
| Johanna Fall | Institute of Marine Research | johanna.fall@hi.no |
| John Simmonds | Unaffiliated | e.j.simmonds1@gmail.com |
| José De Oliveira | Cefas Lowestoft Laboratory | jose.deoliveira@cefas.co.uk |
| Kristiina Hommik | University of Tartu | kristiina.hommik@ut.ee |
| Laura Wise | Portuguese Institute for the Sea and the Atmosphere | lwise@ipma.pt |
| Laurence Kell | SEA++ | laurie@seaplusplus.co.uk |
| Leire Ibaibarriaga | AZTI Sukarrieta | libaibarriaga@azti.es |
| Lies Vansteenbrugge | The Flanders Research Institute for Agriculture, Fisheries and Food | lies.vansteenbrugge@ilvo.vlaanderen.be |
| Marc Taylor | Thuenen Institute | marc.taylor@thuenen.de |
| Martin Pastoors | Pelagic Freezer-Trawler Association | mpastoors@pelagicfish.eu |
| Massimiliano Cardinale (Chair) | SLU Department of Aquatic Resources-SLU Aqua | massimiliano.cardinale@slu.se |
| Michaël Gras | Joint Research Centre | michael.gras@ec.europa.eu |
| Mikael van Deurs | DTU Aqua, National Institute of Aquatic Resources | mvd@aqua.dtu.dk |
| Mikaela Bergenius Nord | SLU Department of Aquatic Resources-SLU Aqua | mikaela.bergenius.nord@slu.se |

| Member | Dept/Institute | Email |
|---|---|---|
| Nicolas Goñi | Natural Resources Institute Finland | nicolas.goni@luke.fi |
| Patrick Lynch | NOAA Fisheries | patrick.lynch@noaa.gov |
| Pekka Jounela | Natural Resources Institute Finland | pekka.jounela@luke.fi |
| Richard D. Methot | Northwest Fisheries Science Center | richard.methot@noaa.gov |
| Rishi Sharma | The Food and Agriculture Organization of the United Nations | Rishi.Sharma@fao.org |
| Santiago Cerviño | Centro Oceanográfico de Vigo | santiago.cervino@ieo.es |
| Sarah Millar | International Council for the Exploration of the Sea | sarah-louise.millar@ices.dk |
| Sindre Vatnehol | Institute of Marine Research | sindre.vatnehol@hi.no |
| Sofie Nimmegeers | The Flanders Research Institute for Agriculture, Fisheries and Food | sofie.nimmegeers@ilvo.vlaanderen.be |
| Sven Stoetera | Thünen-Institute of Baltic Sea Fisheries | sven.stoetera@thuenen.de |
| Tanja Miethe | Marine Laboratory | t.miethe@marlab.ac.uk |
| Tomas Gröhsler | Thünen-Institute of Baltic Sea Fisheries | tomas.groehsler@thuenen.de |
| Valerio Bartolino | SLU Department of Aquatic Resources-SLU Aqua | valerio.bartolino@slu.se |
| Vanessa Trijoulet | DTU Aqua, National Institute of Aquatic Resources | vtri@aqua.dtu.dk |