WORKSHOP OF FISHERIES MANAGEMENT REFERENCE POINTS IN A CHANGING ENVIRONMENT (WKRPCHANGE, OUTPUTS FROM 2020 MEETING)

VOLUME 3 | ISSUE 6

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 3 | Issue 6

WORKSHOP OF FISHERIES MANAGEMENT REFERENCE POINTS IN A CHANGING ENVIRONMENT (WKRPCHANGE, OUTPUTS FROM 2020 MEETING)

Recommended format for purpose of citation:

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

Editors

Jeremy Collie • Daniel Howell • Anna Rindorf

Authors

Eider Andonegi • Per Arneberg • Valerio Bartolino • Aaron M. Berger • Bjarte Bogstad • Lisa Borges • Noel Cadigan • Massimiliano Cardinale • Jeremy Collie • Mikael van Deurs • Sarah Gaichas • Jenni Grossmann • Jan Horbowy • Daniel Howell • Melissa A. Karp • Mathieu Lundy • Wendy E. Morrison • Martin Pastoors • Alfonso Pérez Rodriguez • Jay Peterson • David Reid • Anna Rindorf • Amy Schueller • Henrik Sparholt • Vanessa Trijoulet • John Woodlock • Fan Zhang

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

Contents

i Executive summary

WKRPChange was tasked with examining how ICES handles the estimation of target and limit reference points in the face of changing environmental conditions. In particular, the meeting participants were asked to review the robustness of the current ICES procedures and to suggest specific improvements that could be made, especially relating to changes in stock productivity arising from environmental conditions, species interactions, and density-dependent effects. Part of the work involved reviewing the basis of the ICES reference points and contrasting the ICES procedures with those in the USA and Canada, and part on providing specific guidance for future reference point estimation within ICES.

One common approach to changing environmental conditions is to truncate data series. WKRPChange agreed that this may be necessary in some cases, but several studies were presented showing that the estimation of reference points becomes unreliable (both noisy and potentially biased) as the time series is reduced, and therefore recommended that modelling the specific process involved is generally a better approach than truncation. The meeting noted several examples within the current ICES management system for which reference points are allowed to vary (e.g. F in the case of NEA cod, B_{lim} in the case of Iberian Sardine) according the prevailing conditions. WKRPChange noted that this was only required if conditions were expected to change significantly over the lifespan of the reference points, and that where it was implemented the status determination (the "traffic lights") should be made accordingly.

The key recommendation of WKRPChange is consistent with the conclusions of WKGMSE2, namely that a scoping exercise should be undertaken for each stock to identify any key drivers. Where there is good evidence for ecosystem-driven changes in stock productivity that process should be accounted for in setting reference points. The meeting highlighted that reference points have a finite lifespan, generally related to the benchmark cycle, and the estimation of the reference point should predominantly take into account processes likely to be important over that lifespan. WKRChange noted that many ICES stocks are managed by Harvest Controls which are evaluated through a MSE process. In this case there is considerable scope for including such environmentally driven processes in the Operating Model. However, many stocks are managed through the standard ICES HCR with reference points derived through the EqSim program. There is therefore a specific recommendation that the possibility to include density-dependent growth be incorporated into EqSim, to allow more realism to be included in the estimation of reference points where the evidence indicates that this is important within the reference point life span. The meeting also highlighted the recent work at WKRISH 6, which gave scope to "fine tune" the F_{target} to account for small changes in environmental drivers without requiring full re-estimation of the reference points.

ii Expert group information

Expert group name	Workshop of Fisheries Management Reference Points in a Changing Environment (WKRPChange)
Expert group cycle	Annual
Year cycle started	2020
Reporting year in cycle	1/1
Chairs	Jeremy Collie, USA
	Daniel Howell, Norway
	Anna Rindorf, Denmark
Meeting venue(s) and dates	21–24 September 2020, online meeting, (28 participants)

1 WKRPChange terms of reference and approach

The Workshop of Fisheries Management Reference Points in a Changing Environment (WKRP-Change) met from 21–24 September 2020 by remote means to address the terms of reference to:

a) Review the robustness to environmental and ecological change (e.g. environmental, density dependent or ecological shifts in productivity and distribution) of the current ICES concepts for estimation and application of fisheries target, range and limit reference points.

b) Define appropriate changes in estimation and application of fisheries target, range and limit reference points; within the framework of the management objectives of robustness of advice, precautionarity and yield, in response to:

   i. environmentally induced changes in stock productivity,

   ii. change in species interactions,

   iii. stock density induced changes in stock productivity,

c) Propose a stepwise approach to making appropriate changes in target, range and limit reference points.

The workshop reviewed a series of presentations and publication to incorporate the current understanding and best available science to address the terms of reference in three subgroups. The group was chaired by Anna Rindorf (Denmark), Jeremy Collie (USA) and Daniel Howell (Norway).

The workshop participants defined productivity as the combined effects of recruitment, growth, maturity and natural mortality. Ideally, the selectivity pattern should be investigated as well but this was not addressed in the workshop. Further, there was limited information on maturity and as a result, this was not discussed.

2 Existing principles in reference point estimation in the ICES area

Fisheries management reference points provide standards against which fish stock abundance and fishing pressure are compared. ICES has a well-developed framework for estimating reference points (ICES, 2017a) and for using them to provide management advice (ICES, 2020a). Similar frameworks exist in the US (DiCosimo *et al*., 2010, Methot *et al*., 2014).

2.1 ICES approach to limit, target and range reference points

ICES reference points for analytical assessments follow the precautionary and MSY approach. The two are not mutually exclusive as the MSY approach implicitly includes precautionary approach considerations. Full details of the ICES reference point calculations are available here:

https://doi.org/10.17895/ices.pub.3036

The reference points for analytically assessed stocks (category 1) are absolute. There are very few examples in ICES of relative reference points for analytically assessed stocks (category 2). The reference points are not updated every time a new assessment is run, but at a benchmark process, roughly every five years. The values are then fixed for the time period up to the next benchmark.

Reference points may be re-calculated based on new information or knowledge about the stock/fishery dynamics in a Management Strategy Evaluation (MSE). ICES does not currently have formal guidance about when to calculate and use reference points from an MSE but following a meeting in October 2020 (WKGMSE3) this should be made available.

SUMMARY TABLE OF ICES REFERENCE POINTS			
	Reference Point	Purpose	Basis
PA	B_{lim}	A deterministic biomass limit below which a stock is considered to have reduced reproductive capacity.	The biomass at which recruitment is observed to begin to decline with SSB: the break point of a segmented regression. Or, given knowledge on the history of exploitation, the **biomass that should form a lower limit to exploitation**, such as a lower biomass where high recruitment has been observed.
PA	F_{lim}	An upper limit to exploitation rate, above which exploitation is considered to be unsustainable.	The F that in equilibrium will maintain the stock above B_{lim} with a 50% probability. Or, the **F that will give SSB=Blim** given mean recruitment.
PA	B_{pa}	To provide a stock status reference point above which the stock is considered to have full reproductive capacity.	The value of the estimated SSB that ensures that the true SSB has **less than 5% probability of being below B_{lim}**, i.e. the upper 95 percentile on the distribution of the estimated biomass if the true biomass is at B_{lim}. $B_{pa} = B_{lim} * \exp(1.645 * \sigma)$ where σ is the std of ln(SSB) in final the assessment year
PA	F_{pa}	To provide an exploitation rate reference point below which exploitation is considered to be sustainable, having accounted for estimation uncertainty.	The fishing mortality that, if applied as target in the ICES MSY advice rule (AR) would lead to **SSB≥ Blim with a 95% probability (also know as Fp05)**. The derivation of Fpa should include expected stochastic variability in biology and fishery, as well as advice error.
MSY	F_{MSY}	The F expected to give maximum sustainable yield in the long term.	**F giving maximum yield** given current assessment/advice error, biology and fishery, **constrained so that F_{MSY} is less than $F_{P.05}$** ($=F_{pa}$) when applying the ICES MSY Advice Rule (AR).
MSY	MSY $B_{trigger}$	A **lower bound to the biomass for MSY exploitation**. The point at which F is reduced when applying the ICES MSY AR.	MSY $B_{trigger}$ = Maximum of (B_{pa}, 5 percentile on distribution of SSB when fishing at F_{MSY})
Short-lived	$B_{escapement}$	A deterministic biomass limit below which a stock is considered to have reduced reproductive capacity, including any identified additional biomass need.	**Blim plus an additional biomass** if the advice is based on a deterministic forecast.
Short-lived	F_{cap}	The limit to F that is used when advising catch and probability of SSB>$B_{escapement}$ is not estimated directly.	Based on Stochastic simulation that show a **less than 5% probability of SSB<$B_{escapement}$**

Table 1. Summary of ICES reference points.

The starting point for all the ICES reference points is B_{lim}: the biomass below which recruitment declines with decreasing SSB (Table 1). This can be found either by a segmented regression, where B_{lim} is chosen as the break point; or, given knowledge on the history of exploitation, B_{lim} can be taken as the lowest biomass from which a high recruitment has been observed. F_{lim} is then calculated from B_{lim}, most commonly from a long-term stochastic projection that gives a 50% probability of SSB > B_{lim}. The stochastic projection is based on biological parameters, fishery selectivity in the stock assessment and stochastic recruitment around a segmented regression with breakpoint at B_{lim}.

Next, B_{pa} is calculated using the error from the current stock assessment. B_{pa} is defined as the SSB value such that when SSB is estimated to be at B_{pa}, the probability that the true SSB is greater than B_{lim} is at least 95% in that year. If the error is not available, B_{pa} is estimated as the product of 1.4 and B_{lim}. Previously, the fishing mortality F_{pa} corresponding to a median biomass of B_{pa} (F_{pa}) was estimated. However, from 2020 there is a new ICES definition of F_{pa} as the fishing mortality that, if applied as the target in the ICES MSY advice rule (AR) would lead to SSB being above or equal to B_{lim} with a 95% probability in the long term (also known as $F_{p0.5}$). The derivation of F_{pa} now includes expected stochastic variability in the fishery and biology, as well as advice error.

The ICES MSY advice rule (AR) can be described with a diagram (Figure 1.).

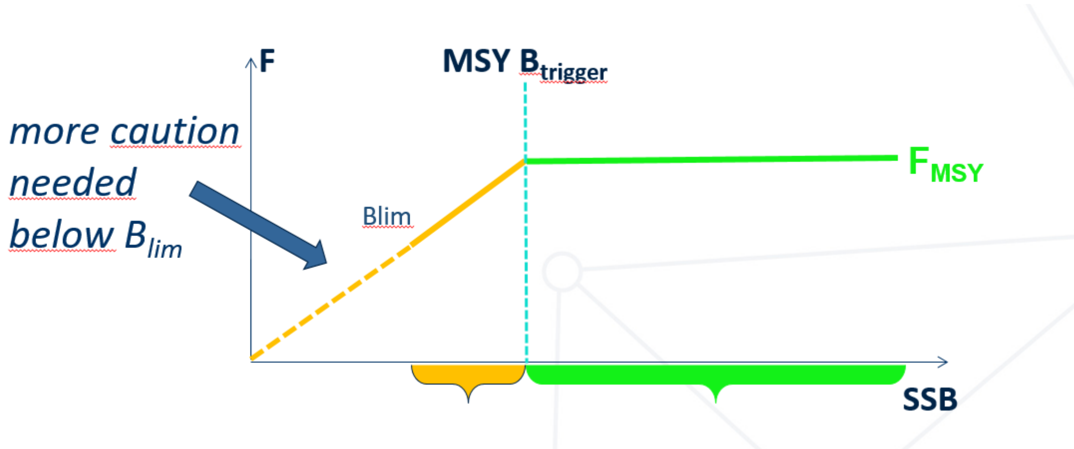

Figure 1. ICES MSY advice rule.

The rule aims to maximise long-term yield whilst safeguarding against low SSB, and uses a combination of F_{MSY} and MSY $B_{trigger}$ to ensure the probability of SSB being below B_{lim} is less than 5%.

If the F following from applying the MSY advice rule is insufficient to bring the stock above B_{lim} in the short term, the ICES advice will be based on bringing the stock above B_{lim} by the end of the next TAC year (short term). This may result in a zero advice. A non-zero advice may be given for stocks that are below B_{lim} if they have an agreed management plan that has been evaluated to be precautionary and to allow recovery in the medium to long term.

The MSY reference points are most commonly calculated with the stochastic simulation software EqSim, although other programs may be used.

In recent years, the EU has requested that ICES provide advice for non-shared stocks using F_{MSY} ranges, where available. The ranges are derived based on yields within 95% of yields at F_{MSY} and as a result deliver less that a 5% reduction in long-term yield compared to MSY. The values around F_{MSY} are based on stochastic simulations, most frequently using EqSim. Should values in the range exceed F_{pa}, the upper part of the range is replaced by F_{pa}, thereby ensuring that all values in the range are in accordance with precautionary principles. The ICES catch advice at

F_{MSY} and at F_{upper} and F_{lower} will follow the ICES MSY advice rule. The entire range is defined to be precautionary.

There is a slightly different process for short-lived species – that is species with a life span of around 4–6 years, high recruitment variability, a high M (1 or greater), highly influenced by the environment and an F generally lower than M. In these stocks, incoming recruitment is often the main or only component of the fishable stock. B_{lim} and B_{pa} are calculated in the same way as for other stocks. The ICES PA and MSY advice for short-lived species is based on biomass escapement strategies. Two approaches exist, one estimates the catch as the amount required to bring the stock to B_{pa} with a requirement that F does not exceed an F_{cap} which has been demonstrated in an MSE to ensure a 95% probability of SSB being above B_{lim}. The other option is to estimate the catch that provides a 5% risk of achieving a biomass below B_{lim}. F_{lim} and F_{pa} are not used for these stocks.

Ongoing work in ICES is looking at stocks for which there is no clear S-R type, to explore whether reference points based on fractions of B_0 or B_{MSY}, similar to the US approach described below, would be more appropriate. There are workshops planned for 2021 following on from WKREBUILD in 2020, a workshop looking at rebuilding plans and how to operate below B_{lim}. Potential workshops on reference points are also envisaged.

2.2 US approach to reference points

US federal fishery objectives and standards for management are established in the Magnuson-Stevens Fishery Conservation and Management Act (DiCosimo *et al.*, 2010; Methot *et al.*, 2014; eCFR 2020). Maximum Sustainable Yield (MSY; or a proxy) establishes the limit fishing mortality rate (F_{MSY}) and the target stock biomass (B_{MSY}). Fishing rates in excess of F_{MSY} represent overfishing, and stocks below ½ B_{MSY} are overfished. Overfished stocks must be rebuilt to B_{MSY} within 10 years (but rebuilding time may be linked to generation time for slow-growing stocks). Eight regional Fishery Management Councils establish fishery management plans to comply with these general rules using the National Standards guidelines on specifying MSY, status determination criteria, control rules incorporating scientific uncertainty, management measures incorporating management uncertainty, and rebuilding plans. National standard guidelines require re-estimation of MSY with "changes in long term environmental or ecological conditions, fishery technological characteristics, or new scientific information." Guidelines state that species interactions should be taken into account in specifying MSY, if practical, or considered when setting optimum yield below MSY if not considered in the specification. Reference points must be re-specified "if environmental, ecosystem, or habitat changes affect the long-term reproductive potential of the stock…." and are revised on a research track schedule of 5–10 years.

All Councils derive reference points and quotas from stock assessments using a process that reduces from an overfishing limit (OFL; catch from fishing at F_{MSY}) to an acceptable biological catch (ABC) incorporating scientific uncertainty, to an annual catch limit (ACL) including accountability measures for the fishery, and annual catch target (ACT) incorporating management uncertainty. There is variation across regional Councils in how the guidance is applied to set MSY proxy reference points. Harvest control rules also differ across Councils and sometimes between FMPs within the same Council, but many include a fixed F rate above the target B (B_{MSY}), with F sloping down below target B to reduce fishing pressure (and to avoid stocks dropping below the overfished threshold, ½ B_{MSY}). Reference points are estimated within stock assessment development and review processes, using similar methods to those of ICES (described above). Consideration (or not) of environmental change, species interactions, density dependence, and other factors in determining reference points generally happens within individual stock assessment working groups and reviews. Stock assessment working groups can spend considerable effort in deciding how much of a stock's historical data to include in reference point calculations, balancing the stability from longer time series with the need to represent "prevailing ecological, environmental, and fishery technological characteristics" in MSY.

2.3 Canadian approach to biological reference points

In Canada, the identification and application of biological reference points are based on a fishery decision-making framework incorporating the Precautionary Approach (PA) (DFO, 2009). The core is a decision framework (Figure 2), which classifies the status of a fish stock into three zones along the abundance or biomass axis according to its Limit Reference Point (LRP) and Upper Stock Reference (USR), i.e., LRP lies on the boundary between critical and cautious zones and USR divides cautious and healthy zones. Stocks are also classified into two "fishery status" zones based on whether fishing mortality is at or below, or exceeds, its Removal Reference (RR; a maximum or limit reference point for fishing mortality rate). The LRP represents stock abundance or biomass levels below which the stock is considered to be at risk of serious harm, typically interpreted as reproductive impairment; management actions are intended to avoid LRP breaches. The USR can represent either a target or threshold reference point, and is positioned by fisheries managers to represent a desired stock abundance level and/or a level at which management measures must change to avoid an LRP breach, while taking into account both conservation and socio-economic considerations for the fishery. Management measures are generally intended to maintain stocks at, or above, the USR, in the healthy zone. The PA Policy of Canada generally considers reference points to also represent operational control points, in that its default guidance provides for different management priorities and changing harvest strategies across the three zones. In the critical zone, priority is given to conservation considerations and removals from all sources must be kept at the lowest possible level until the stock has rebuilt above the LRP. In the cautious zone, conservation and socio-economic considerations should be balanced and removals from all sources should be set to promote fish stocks to grow to healthy zone. In the healthy zone, priority is given to socio-economic considerations and removals from all sources should be kept below the RR. The Target Reference Point (TRP), if one is set, should be set as equal to or above USR.

Under this fishery decision-making framework, the explicit methods to calculate biological reference points are fairly diverse and tend to be case-specific. This framework is designed to be less prescriptive and give much freedom to scientists, managers and stakeholders to identify biological reference points based on the ecological, social and economic context of specific fish stocks. As for guidance purposes, some recommendations in the PA Policy are given as best practices when more suitable stock-specific reference points have not been identified. For example, the LRP and USR may be set as 40% and 80% of B_{MSY}, respectively. The RR, however, must be less than or equal to F_{MSY}. For specific stocks, other methods can also be used to identify LRP, USR and the RR, as long as they are consistent with intent of the fishery decision-making framework.

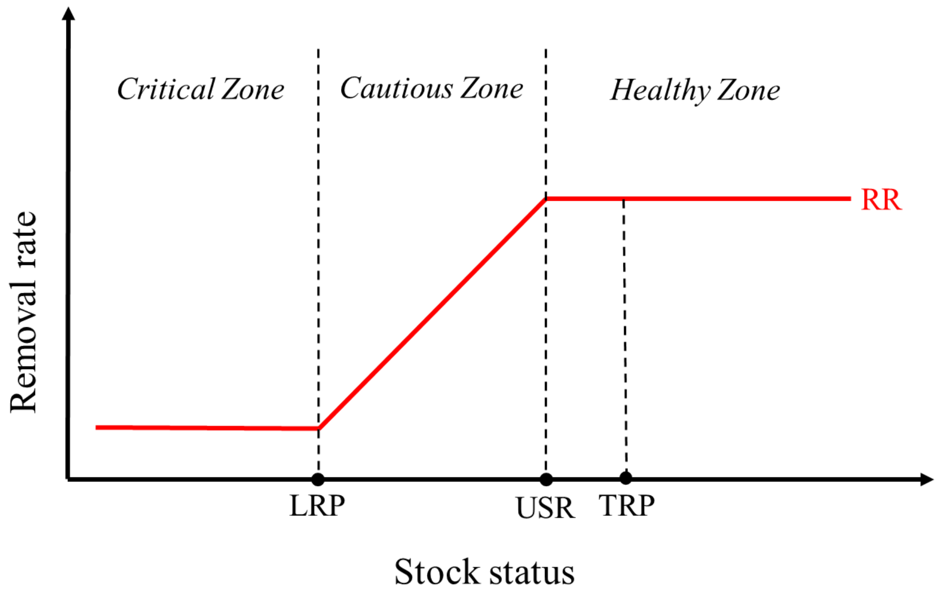

Figure. 2. The three-zoned decision framework. LPR is limit reference point, USR is upper stock reference, and TRP is target reference point. The red line is the removal reference (RR) among three zones.

2.4 Bias in stock-recruitment relationship based biomass limit reference points

Using a combination of data simulations and data from 51 small-bodied pelagic fish stocks, van Deurs *et al*. (2020) analysed the sensitivity of B_{lim} to choice of method (type-1 or type-2, see Figure 1 in van Deurs *et al*.), time-series length, and stock development (e.g. rebuilding or declining). The study investigated two versions of type-1; good recruitment defined as being above the 50th percentile (P0.5) and 80th percentile (P0.8), respectively. The type-2 version used the grid-search method to find the optimal breakpoint of a segmented regression (Hockey stick, HS) (recommended by Barrowman and Myers, 2000). When recruitment variation is low, the HS method seems promising (little effect of time-series length and high precision). It should be noted that choosing other methods to identify the breakpoint may yield very different results. However, variation in recruitment can be substantial, particularly in small-bodied pelagic fish stocks. Hence, it may be useful to consider other methods in addition to HS. P0.8 produces on average the same B_{lim} as HS, but the correlation between HS and P0.8 on a stock-by-stock basis is surprisingly low. P0.5 is relatively precise, but systematically underestimates HS. Both P0.5 and P0.8 are sensitive to time-series length as B_{lim} decreases with increasing time series length unless the stock size is increasing substantially over time. In general, decreasing stock trends cause B_{lim} to decrease over time as more years are added. Lastly, selected stocks were used to demonstrate that B_{lim} is associated with substantial uncertainty and that the different methods behave differently from stock to stock (one size does not fit all).

The conclusion is that shortening time-series (e.g. to account for regime-shifts), particularly in stocks increasing or decreasing trends in stock size (e.g. caused by trends in productivity affects the trend in B_{lim} over time) may substantially affect results. Due to the high variability in the stock-recruitment relationship, having a shorter time-series may also substantially increase the CV of the estimated reference point.

2.5 Time series length effect on F_{MSY} estimation

Sparholt *et al*. (2020) conducted a sensitivity analysis using the Faroese stocks of cod, haddock and saithe to investigate regime shifts and their influence on stochastic production model estimates of F_{MSY}. A 39-year time-window was moved in steps of 10 years, and they applied the stochastic production model (SPM) to estimate F_{MSY} for each time-window.

The time series analyses showed that F_{MSY} (expressed as catch/biomass at maximum productivity) in the past century has been relatively stable for cod (Figure 3). For haddock and saithe, it has fluctuated by a factor of about 2.

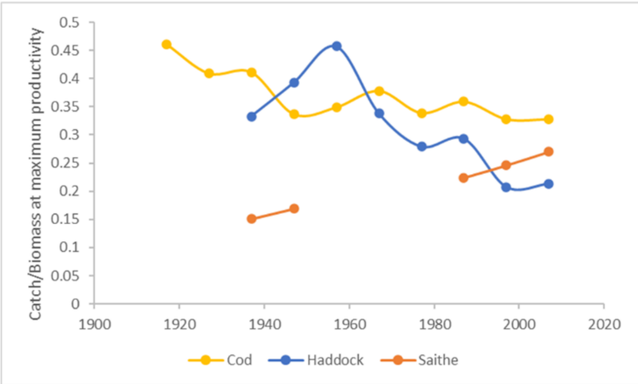

Figure 3. Estimates of catch divided by exploitable biomass at maximum productivity (B_{MSY}) of cod, haddock, and saithe in Faroese waters (i.e. F_{MSY}). Based on SPMs applied for intervals of 39 years that were moved in steps of 10 years represented by the dots. Smoother curves were applied between the dots.

SPiCT was used to test the sensitivity of the estimated F_{MSY} to the length of the time series (Sparholt *et al*., 2020). Catch/F was used as the biomass index. The F_{MSY} estimate is constant until the time series used begins in or later than 1985, at which point the estimate drops and the uncertainty range expands (Figure 4). The analysis indicates the importance of long time series in estimating reference points.

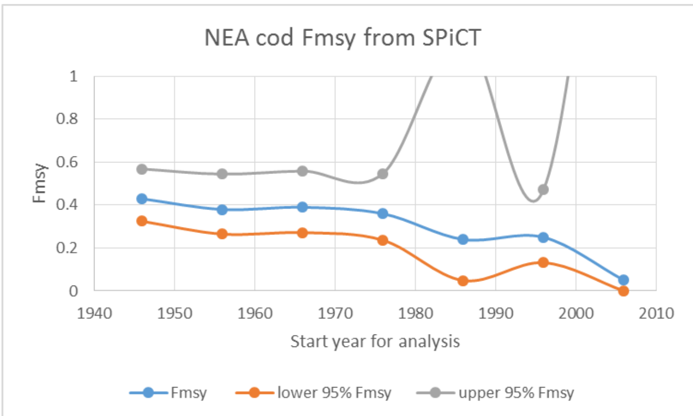

Figure 4. F_{MSY} as a function of start year in the time series. Analysis used data from the year on the axis up to 2015, thus points on the right have less data in their estimation. The 95% confidence intervals are also shown.

2.6 Differences and similarities among approaches

The approaches in the three areas differ in three aspects: the definition of biomass limits to stock biomass, the placement of precautionary buffers and the role of aspects other than yield and recruitment in the setting of reference points.

The ICES system bases biomass limits reference points on the stock recruitment relationship, which makes the definition of breakpoints a key concern in this area. In contrast, the US approach bases biomass limit reference points on B_{MSY}, and hence integrates impacts of both growth overfishing (fishing at an F higher than that providing maximum yield per recruit) and recruitment overfishing (fishing at an F higher than that affecting recruitment). This means that a decrease in yield due to e.g. recruitment overfishing can be negated through an increase in yield per recruit. This can occur in species with high natural mortality or density dependent growth, where maximum yield per recruit tends to occur at high F levels. Both methods are sensitive to the shape of and variability around the stock recruitment relationship. The ICES method is sensitive to the method used to estimate the stock recruitment breakpoint whereas the US method can be based on simulations and hence does not require breakpoint analyses. The biomass limits estimated differ between the two methods:

- If the US estimates of B_{MSY} do not include a stock-recruitment relationship, the two methods will differ substantially.
- If there is no stock-recruitment relationship (random variation with no sign of decrease at low stock size), the ICES approach uses lowest observed value as a proxy for either B_{lim} or B_{pa}. In this case, F_{MSY} is equal to the F providing the maximum yield per recruit but the ICES biomass limit reference point to be avoided with a 95% probability is not necessarily related to B_{MSY}, particularly for stocks that have historically been fished far from F_{MSY} (in both directions). Hence, in these cases, the ICES approach will lead to different biomass limits than the US approach (higher or lower depending on the stock history).
- If the stock-recruitment relationship is proportional or close to this for most of the observed biomass range, both the US and ICES methods will tend to estimate the reference point at high biomass levels compared to those historically observed: the ICES approach because the HS breakpoint will be estimated in the upper part of the biomass range and the US approach because B_{MSY} will tend to be estimated well above the observed biomass range. In the ICES approach, there is an option to define the biomass reference point by expert judgement if fishing mortality has historically been moderate and hence lower biomass limits can also be estimated. Hence, it is likely that these cases can lead to lower biomass limits using the ICES approach than when using the US approach.
- If the stock endures a high natural mortality or density dependent growth, F_{MSY} may be estimated at high levels and B_{MSY} as a result will be low.
- If there is a stock recruitment relationship with a decrease in recruitment at high stock size (Ricker type), B_{MSY} tends to be estimated around the maximum recruitment, whereas the ICES approach tends to estimate B_{lim} at the biomass providing 83% of the maximum recruitment. Hence, there is no guarantee that the two are comparable, and it would seem likely that the US biomass limit ($0.5B_{MSY}$) is lower than the ICES limit.
- In the special case where recruitment decreases below $0.5B_{MSY}$ the estimates from the two approaches will be identical. However, this is likely to occur infrequently and generally, the two approaches will provide different biomass limits. In general, there is no guarantee that B_{MSY} exceeds the ICES biomass limit, let alone that it is twice the value (Rindorf *et al.*, 2017).

The second difference between the US and ICES approaches is the placement of precautionary buffers. In ICES, the precautionary buffer is applied to the biomass of the stock (B_{pa}) and the TAC is estimated from F_{MSY} except if the stock is expected to decrease below B_{pa}. In the US approach, the precautionary buffer is applied to ensure that F remains below F_{MSY}.

The Canadian approach is most flexible and can potentially mimic either the US or ICES methods depending on the reference points chosen for a particular specific stock.

2.7 Conclusions on the robustness to environmental and ecological change of the current ICES concepts for estimation and application of fisheries target, range and limit reference points.

Generally, the reference point is considered a fixed standard against which the stock fluctuates. However, there is increasing awareness that the reference points themselves will vary with changes in demographic parameters, species interactions, and other environmental changes.

The estimation of MSY reference points includes fitting the functional form of a stock-recruitment relationship. Attention is therefore focused on model choice and uncertainty in the parameter estimates. Both the choice of model and number of years included in the fit can strongly affect the estimated reference points. Work presented at the meeting suggested that the breakpoint of a hockey-stick recruitment function may be appropriate for estimating B_{lim} where there is low variability in recruitment and the breakpoint thus well estimated. Where the breakpoint is poorly estimated then measures such as the biomass corresponding to the 50 or 80th percentile of recruitment may be a more useful estimator. In this case the choice of percentile reflects the degree of risk aversion of the decision makers.

In general, WKRPChange considered that the ICES process for estimating and applying reference points is able to account for many of the ecosystem induced uncertainties through the iterative re-evaluation of reference points through periodic benchmarks or MSE exercises and the inter-benchmark process to account for urgent revisions where conditions require. It is important however to ensure that the estimation includes the relevant fisheries and biological processes. Critically, identifying trends in variance (including autocorrelation) and trends in mean values are needed for adequately estimating reference points. While truncating time series is one approach to dealing with changing conditions, this needs to be used with caution as reduced length of time series can lead to errors in estimating reference points.

3 Environmentally induced changes in stock productivity (ToR (b)i)

Environmental variability affects fish stocks over a range of time scales: short (interannual), medium (5–10 years), and long (decadal). It is important to differentiate the patterns of variation (e.g. high-frequency white-noise variation versus low-frequency regime-like variation) because the effects on stock productivity and corresponding reference points depend on the time scale. Short-term variability is included as random deviations in stochastic simulations.

Medium-term variability occurs on the same timescale as fish lifespans and the management system. This variability can be accommodated by regular updates of reference points (e.g. 5 yr). For some stocks, the environmental drivers and processes by which they affect fish production are known. A pertinent example concerns the ecosystem changes in the Irish Sea and the management measures that have been introduced in response to these changes.

3.1 Productivity changes with known drivers: ecosystem changes in the Irish Sea

The fisheries of the Irish Sea have changed dramatically over the last 50 years. At the start of this period, the fishery was finfish dominated, consisting predominantly of cod, whiting, and Atlantic herring. Since 1970, landings have declined by 97% for cod, by 88% for whiting, and by 81% for Atlantic herring. Over the same period, invertebrate landings increased, mainly composed of the Norway lobster *Nephrops norvegicus*, (+56%), crabs (+78%), and scallops (+34%). Landings of *Nephrops* remained relatively stable from the 1980s to the mid-2000s, with some declines evident since then (ICES, 2019a). From the early 2000s, management measures were implemented, reducing fishing effort by around 90% since 2003 for both the whitefish otter trawlers and the beam trawlers. Effort by the trawlers targeting *Nephrops* declined by around 30%. The main purpose of this management approach was to improve the cod stock, but it was unsuccessful; there has been no recovery to B_{MSY}. Additionally, whiting has failed to recover to previous stock levels. For other stocks, namely plaice, haddock, and Atlantic herring, there has been stock recovery since the early 2000s, likely helped by the substantial effort reductions.

In 2014, the North Western Waters Advisory Council (NWWAC; an EU mandated fisheries stakeholder forum for both industry and environmental groups) asked ICES to investigate why the substantial effort reductions had not helped with recovery of cod, whiting, and sole, and if the lack of recovery could be linked to environmental factors. Based on this request, ICES set up a benchmark workshop series (WKIrish: ICES, 2015) to examine the single-species stock assessments and the possibility of ecosystem drivers having a role in the changes. The first part of the work involved a wide-ranging scoping workshop involving scientists and fishery stakeholders, which was followed by an assessment benchmark process developing and improving the single-species stock assessment data and methodology, as many of the assessments were unreliable (ICES, 2017b, 2017c). Later, the workshops focused on developing a suite of ecosystem models in explicit collaboration with the NWWAC stakeholders for the Irish Sea. These modelling approaches included LeMans (Thorpe and De Oliveira, 2019), MoSES (ICES, 2020b), and an EwE model (Bentley *et al*., 2020). The LeMans and MoSES models were not fully operational in time for the final workshop, but the EwE model was available for operational advice, and thus used.

The EwE model included 41 functional groups, including the commercial species as adults and juveniles, as well as other groups ranging from detritus, discards, and primary producers to mammals and seabirds (Bentley *et al*., 2020). The different commercial fleets were included with their effort, as well as temperature, top-down (e.g. predation) and bottom-up (e.g. primary production) interactions, and the North Atlantic Oscillation (NAO) anomaly, all of which were identified as significant drivers of historic biomass and catch trends. The interdisciplinary approach combined the expertise of three types of experts: ecological modellers, stock assessors, and stakeholders. A key element of the work was the continuous involvement of the stakeholders (both industry and environmental bodies) who provided pivotal information for the diets of many key species in the model, particularly for 1973, the start year for the model (Bentley *et al*., 2019a). They also provided critical information on effort trends by gear, starting well before formal records that begin in 2003 (Bentley *et al*., 2019b).

Ecosystem drivers were identified through a hypothesis testing process that evaluated goodness of fit with and without ecosystem drivers included as forcing functions in the Ecosim model. Significant ecosystem drivers for stock production were identified for four species (ICES, 2020b). Both cod and whiting were strongly influenced by sea surface temperature with a 3-year lag, thus linking to recruitment. Atlantic herring had a strong link to large zooplankton abundance. *Nephrops* were linked to the abundance of predators at trophic level 4 and above. For sole, plaice, and haddock, no convincing ecological indicators (i.e. possessing both strong correlation and mechanism of effect) were identified.

The end products were recommendations for target Fs within the "pretty-good-yield" ranges that have been adopted for many stocks in the EU. ICES provides precautionary F_{MSY} ranges (F_{MSY} upper and F_{MSY} lower) that are derived to deliver no more than a 5% reduction in long-term yield compared with MSY for selected stocks (Hilborn 2010; ICES, 2014, 2019b; Rindorf *et al*., 2017). Using the identified indicator for each stock, the F_{target} value was scaled linearly within the F_{MSY} range according to the current value of the indicator within the historical range during the model tuning period. For example, single species F_{MSY} and associated quotas were adjusted for cod and whiting based on sea surface temperature, Herring F and quotas were adjusted according to empirical estimates of zooplankton abundance, and *Nephrops* F and quota was adjusted based on combined biomass of predators. This adjustment allows the ecosystem understanding to be incorporated within the existing single-stock management framework, and critically, within the F_{MSY} ranges that have already been identified as being precautionary. On this basis, ecosystem information can be used to set F within those ranges, and within the management advice paradigm.

3.2 Productivity changes with unknown drivers

In cases for which the drivers have not been identified, state-space models can be used to track hidden processes over time. State-space models are useful and flexible tools to quantify historical variation in stock-recruitment relationships. Likelihood tests are used to determine whether a stationary or time-varying stock-recruitment model best explains the historical data (Tableau *et al*., 2019). Time-varying S-R models can track changes in per capita recruitment without the need to choose time windows of recruitment data (Minto *et al*., 2013; Zhang *et al*., 2018). Although recruitment may vary with a higher frequency, the underlying stock-recruitment parameters vary more slowly. These more gradual changes imply dynamic reference points that are more in line with the management system. Long time series are needed to reliably estimate the process and measurement error variances (or their ratio) with confidence.

Despite the recognition of environmentally induced shifts in productivity, in few cases has this knowledge been translated to the estimation of reference points and stock status. Berger (2019) compared static and dynamic spawning potential ratios of 18 west coast groundfish species. For the two most variable species, the dynamic ratio implied very different levels of stock depletion.

The MSY proxy reference levels for Georges Bank Atlantic cod were found to be sensitive to observed changes in growth and maturity rates, with the effect of variable growth predominating (Miller *et al.*, 2018). Interestingly, the effect of temperature on growth reversed sign when a hidden autocorrelated process was included in the state-space model. An example is provided for Georges Bank winter flounder, a stock for which there is significant evidence of a time-varying stock-recruitment model (Figure 5). Between 1982 and 2014 productivity declined with a quasi-cyclic pattern. The corresponding estimates of B_{MSY} and MSY scale with productivity. These changes in productivity could be used to periodically update the MSY-based biological reference points.

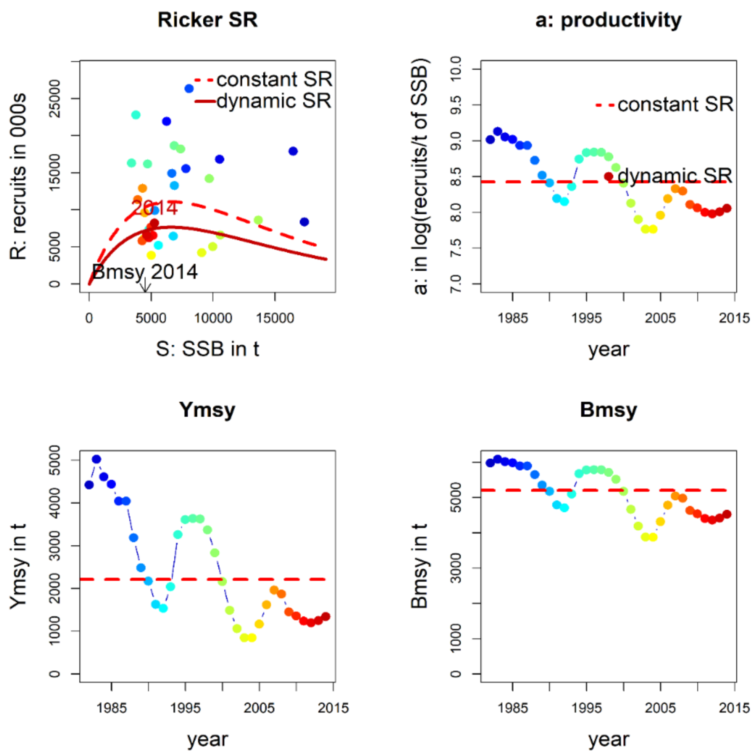

Figure 5. Dynamic stock-recruitment (SR) model of Georges Bank winter flounder (*Psuedopleuronectes americanus*).

Projecting future recruitment with stock-recruitment models is critical to calculate reference points and implement harvest levels. Auto-correlated recruitment projections produce recruitment dynamics that are more similar to historical patterns of variation than stochastic (white noise or resampling) projections (Zhang *et al.*, 2020). Additionally, stochastic projections tend to produce larger values of MSY and F_{MSY} than auto-correlated projections in dynamic ecosystems exhibiting low-frequency and large-magnitude variations (Zhang *et al.*, 2020). Therefore, it is important to match the historical auto-correlation structure when projecting stock-recruitment relationships.

Long-term (decadal) variability may manifest itself as abrupt regime shifts instead of gradual changes. Hidden Markov stock-recruitment models can be applied in the case of abrupt regime shifts. These models estimate separate stock-recruitment parameters for each regime and the probability of shifting between regimes (Zhang, work in prep). An example of an environmentally induced regime was provided for the Iberian sardine.

3.2.1 Regime shifts: the case of the Iberian sardine

Given the low recruitment events for Iberian sardine (lowest historical value in 2017) and the low biomass estimated for the population of age 1 and older during the last decade, ICES advised a zero catch for 2019 for Iberian sardine (ICES 8.c and 9.a), based on a MSY approach. An official TAC is not specified for this stock and the management plan of 2013 was found to be not precautionary in 2017 (ICES, 2017d). A new management and recovery plan is still under development by Spain and Portugal.

As a consequence of these events, ICES received a Special Request from Spain and Portugal in 2019, asking, among other things, for the re-examination of the Biological Reference Points (BRPs) for the Iberian sardine stock, under the assumption that the low productivity of the stock during the last decade, might continue in the future. All this was addressed in an ICES workshop, (ICES, 2019b), following the methodology proposed in ICES (2017a) guidelines for fisheries management reference points. ICES (2017a) assumed two different regimes with different stock productivities: 1993–2017 (medium stock productivity) and 2006–2017 (low stock productivity). The Hockey-stick S-R relationship was adopted for the calculation of reference points in both regimes, with the candidate B_{lim} estimated as the change point of the Hockey-stick model fitted to the data. The biological parameters were similar to the ones used in the last assessment and revised in the last benchmark (ICES, 2017d).

Results showed that, in the medium productivity regime (1993-2017), the candidate B_{lim} (361 639 tonnes) would be 7.2% higher than the value of B_{lim} (337 448 tonnes) used at that moment (so called 'current'). However, in the low productivity regime, the candidate for B_{lim} is 196 334 tonnes (46% lower than the current B_{lim}). This new B_{lim} value calculated for the low productivity regime (B_{lim} = 196 334 tonnes) was adopted by ICES, and the new estimates for B_{pa}, F_{lim}, F_{pa} and F_{MSY} were provided:

Reference point	Previous	Updated
B_{lim} (tonnes)	337448	196334
B_{pa} (tonnes)	574066	252523
F_{lim}	0.250	0.156
F_{pa}	0.189	0.118
F_{MSY}*	0.12	0.032

The two regimes were inferred from stock abundance but mechanisms giving rise to the productivity regimes have not been identified. Adjusting B_{lim} downward in the low productivity regime may be considered risky when the cause of the regime shift and its duration are unknown. Due to the difficulties in predicting the persistence of the current stage of low productivity, ICES recommended that it needs to be monitored regularly, to determine if the BRPs and the resulting HCRs remain valid. Since this evaluation, high recruitment has been observed, challenging the interpretation of regime shifts.

3.3 Incorporating variation caused by known drivers into fishing mortality targets

The F_{eco} approach was presented as a method by which ecosystem information, and the output of ecosystem modelling, can be included into the advice rule, while remaining precautionary and without requiring a major revision of the current advice-giving process. The workflow involved in this method is summarized in Figure 6 below. An ecosystem model is developed and benchmarked alongside the single-species assessment model. All the steps involved in the stock assessment and advice process remain as currently conducted within the realm of the single-species model. The only alteration is that the target F is adjusted, either on an annual or multiannual basis, according to the ecosystem information. A constraint is imposed that the adjustment of the F target cannot exceed the precautionary fishing pressure (F_{pa}) reference point, thus ensuring that the adjusted quota remains precautionary.

Two different examples were presented which had independently adopted a version of this approach, with different management goals. In the US, there was work on Atlantic menhaden, aiming to identify potential fishing reductions on the forage fish due to predator requirements (Section 4.2.1), while in the Irish Sea the aim was to fine tune the fishing pressure to account for environmental variations (Section 3.1). An important step in the process is the determination of the ecosystem indicator to be used to adjust F. Two different approaches were taken in these case studies. For the Irish Sea, the goal was to determine potential environmental drivers of managed fish stocks so that harvest rates could be adjusted accordingly. The ecosystem indicators identified for commercial stocks in the Irish Sea included sea surface temperature, zooplankton abundance, and predator abundance. For Atlantic menhaden, the ecosystem indicator was quite different. Here, menhaden biomass was treated as an ecosystem indicator of food availability to their most sensitive predator, striped bass. Both case studies are described in more detail in this report.

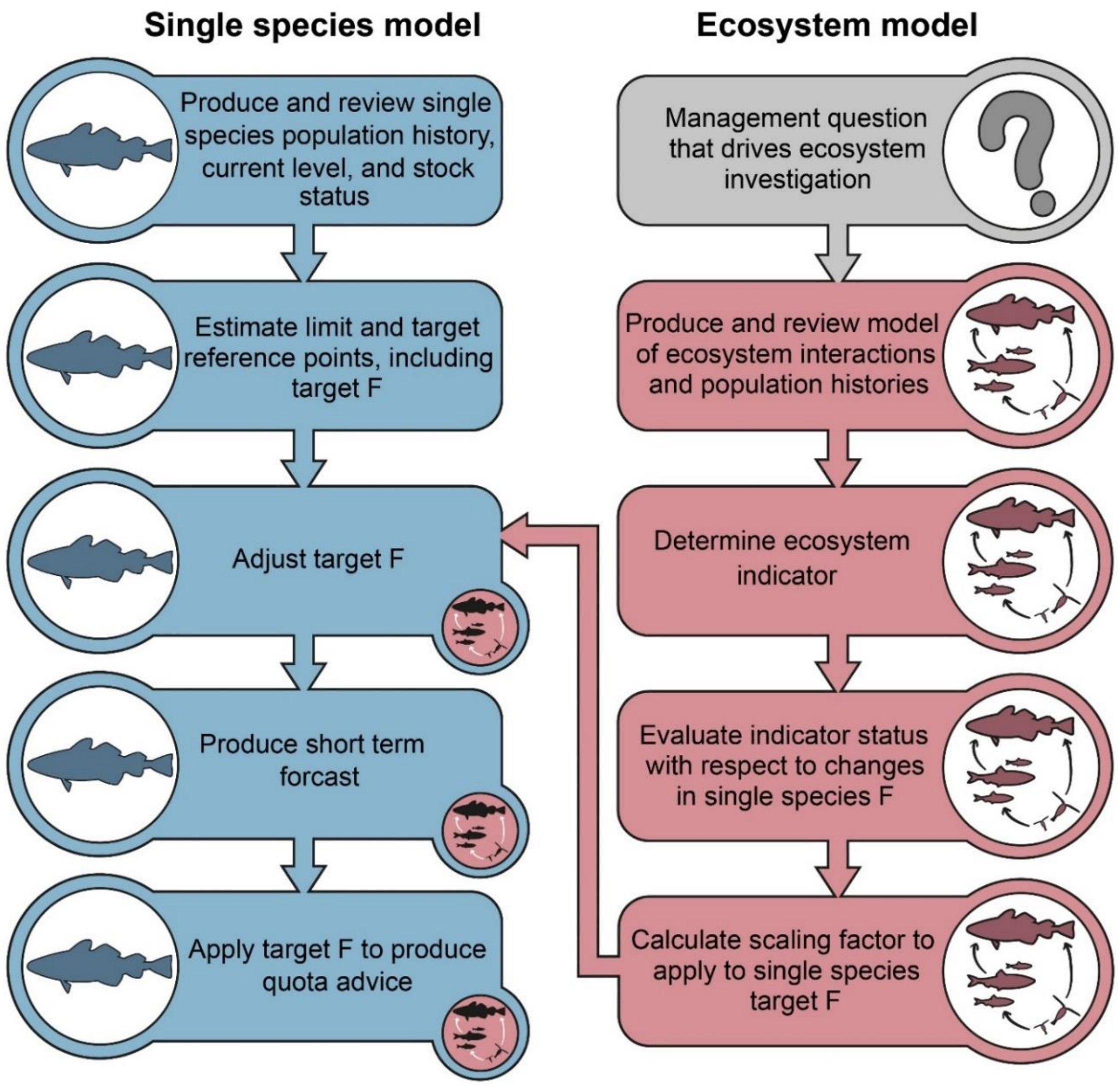

Figure 6. Flow chart outlining the steps in advice giving involved in the proposed method, with the input of the ecosystem modelling to the single-species advice highlighted.

The method of adjusting the F_{target} is not specified in Figure 6, because as the examples above demonstrate the adjustment will depend on the needs of the particular case. Note that this method does not involve directly transferring a value of F_{target} from the ecosystem model to the single-species model. Rather, it applies a scaling factor to the F_{target} from the single-species model, thus ensuring that the resulting F_{target} is compatible with the single-species F reference points from the assessment. In the Irish Sea, F_{eco} was specific for each species and based on a linear scalar to the environmental driver, constrained within defined targets and limits. In the menhaden example, F_{eco} was established based on the response of a single predator.

Finally, the assessment model must then run a short-term forecast (typically 1–3 years) with the revised F_{target} to produce quota advice for the coming year or years. The proposed method adjusts the F_{target} based on ecosystem information as can be seen in the examples described above. The method does not otherwise alter the existing harvest control rule methodology, and the revised F_{target} is constrained to not breach precautionary guidelines.

The strength of this approach is that the assessment and management of fish stocks remains with the single species assessment models and within the current management structure as much as possible. The stock history, status, reference point calculation, initial estimate of target *F*, and the translation of the final target into quota advice all remain within the realm of the single-species assessment model. As a result, the advice system is familiar to managers and stakeholders, so these *first steps* toward EBFM are less onerous and daunting than what managers, stakeholders, and scientists might have expected. Only the adjustment of the target *F* is influenced by the ecosystem modelling, which also implies that no change is required to the existing assessment model. This draws on the strengths of existing single-species stock assessments, while broadening the management approach to include ecosystem considerations. Essentially, the traditional single stock assessment recognizes that stock status can change in response to fishing mortality, and intrinsic population dynamics. The ecosystem model recognizes that stock status can change in response to extrinsic ecosystem factors. The proposed approach then modulates the target *F* from the stock assessment model with the status of the indicator(s) identified by the ecosystem model. In addition, the management framework currently in place can remain and the method simply fits within the existing structures. The management regime itself can, of course, continue to evolve and can change with time as the needs arise, but no fundamental change is required before implementing this method. The proposed method allows the key driver(s) for a particular stock to be considered without requiring that all possible drivers be included. Finally, because the method adjusts F_{target} to produce an F_{eco}, with the constraint that F_{eco} remains at or below existing F_{pa} levels, the risk of stock collapse is no higher, and potentially lower, than under current single species assessments and management.

Apart from the step involving adjusting the F_{target}, this workflow is exactly as is currently done for single species management advice; and thus, offers an easy and straight-forward transition to implementing EBFM. Clearly, the F_{eco} approach does not address every possible issue under EBFM, and it remains preferable to incorporate the relevant processes into the standard simulations where possible. Nevertheless, the method here represents a valuable and flexible step forward in practical, operational EBFM, which can be directly implemented within existing management frameworks.

3.4 Conclusions on changes in reference points as a result of environmental change

Biological reference points are not fixed in time, but are generally revised on a time frame of c. 5 years (the ICES benchmark cycle or MSE cycle). Therefore, they need only be robust for this medium-term timespan to serve their purpose. Where there are no systematic changes in the environment, the "base case method" used by ICES is likely to be robust. However, where there are systematic changes, more detailed investigations are required. The workshop participants agreed that consistent evidence of environmental change is needed before adjusting reference points. This evidence can be based on trends in the mean state as well as patterns in the residuals. Once evidence is obtained, time series may be modelled using additional information, truncated where no additional information exists, or additional processes (e.g. weight at age) incorporated into the reference point calculations.

Truncating time series of weight at age or selectivity is not expected to have as profound effects on reference points as recruitment time series as the variability in these inputs are relatively low compared to the variability in recruitment. If substantial changes have been observed in the last years, truncating the weight at age or selectivity to the most recent period might be appropriate when there are reasons to believe that these changes are not going to be reversed in the period up to next benchmark. This truncation could be especially important for short lived species. In this case, running simulations using the entire time series of weight at age or other factors affecting productivity, and estimating alternative reference points, could avoid an inter-benchmark exercise if changes are reversed.

When truncating the stock and recruitment pairs, one of the considerations is to assess the impacts of trying to estimate a S-R relationship on a smaller number of data points, whereby certain individual data points could have a much larger impacts on the parameter estimates. An example was provided when estimating reference points for North Sea herring (ICES, 2018a). The perception of a regime shift between 1993 and 1994 was based on a paper by Clausen *et al*. (2017). However, when the recent productivity regime was used in the EqSim reference point framework, the resulting in reference points were far outside of the expected range. A sensitivity analysis (Pastoors 2018, in ICES, 2018) demonstrated that this was caused by individual S-R pairs that had a large influence on the shape of the S-R relationships.

Simulation studies have shown that too short a time window can lead to unreliable estimates of reference points (van Deurs *et al*., 2020). There is therefore a trade-off between higher precision gained by using the full time series and higher accuracy obtained by adjusting the time range used to account for regime changes. By truncating time series, one is making the strong assumption that historical data carry no information for the current and foreseeable future. This issue has been studied previously, and recommendations from DFO Canada (DFO, 2013) suggest that truncation should only be considered if it is considered unlikely that the change will reverse in the short-to-medium term, either because there is a mechanistic understanding of the reasons for the regime change or that the current conditions have persisted for an extended period of time. Truncation should not be a "cherry picking" exercise, but should be based on empirical evidence, preferably over multiple stocks in a region, and should make the least change required to fit that evidence.

Truncating the data is a rather crude approach, and ideally one would find a method for accounting for changing environment without having to drastically curtail the time series. This could be done by trying to model bimodality and autocorrelation based on empirical evidence (e.g. Horse Mackerel, de Oliveira *et al*., 2013) or by modelling the mechanism behind the environmental change. To some extent this is done for changing predation mortality in some ICES stocks (e.g. NEA cod, capelin), but could be extended to more processes. For shorter term fluctuations, modelling the variability as an autocorrelated process is a viable approach. State space models are widely used within ICES, and these provide a suitable framework to model hidden processes.

The method for projecting future recruitment (e.g. with and without autocorrelation) can impact the estimation of reference points such as MSY, $F_{MSY,}$ and B_{MSY} (Zhang *et al*., 2020), so it is important to keep and replicate historical patterns in the forecasts used for reference point estimation. Especially, autocorrelation in stock productivity (e.g. runs of good or bad recruitment) affects the F_{MSY} estimates with stronger autocorrelation leading to lower F_{MSY} (Zhang *et al*., 2020). Hence, lighter fishing pressure is needed when there are strong autocorrelations of recruitment, and ignoring autocorrelation where it exists can result in overexploitation. When running simulations, care is needed to ensure that the processes in the simulations remain within the range of the tuning data (both the observed range of values, and the autocorrelation structure). This is an issue with simulation of recruitment in current MSE exercises, but would become more of a concern as more processes are modelled.

4 Change in species interactions (ToR b(ii))

The group reviewed examples of reference points that are conditional on predator or prey abundance either based on historic information or dynamic forecasts.

4.1 Accounting for historic removals by predators in the setting of reference points

4.1.1 Capelin advice dependent on cod in the Barents Sea

Capelin is a short-lived (3–5 years) semelparous pelagic fish, with large natural fluctuations in abundance. It is the main forage fish in the Barents Sea and also the main prey item for cod. Intensive predation by cod takes place in the pre-spawning period (January–March). The fishery is only on maturing capelin in January–March.

Capelin is surveyed by an annual acoustic survey in August/September from which estimates of age 1+ abundance are used as absolute values in the assessment. Maturation is assumed to be length-dependent -— all fish >14 cm in autumn are assumed to mature and spawn in the following year. Spawning stock size is predicted from October 1 to April 1 (spawning time) using historical natural mortality (M) values in October-December and M values estimated using a long time series of cod stomach content data for the period January–March. This means that the stock recruitment relationship is based on temporally variable estimates of natural mortality. The harvest control rule says that the TAC should be set so that there is 95% probability for SSB > B_{lim} (200 kt). The B_{lim} value of 200 kt is somewhat arbitrary, taking into account that the smallest SSB which has given rise to a strong year class is around 100 kt (1989-year class). Since capelin spawns only once and there is no reliable survey on the spawning stock, modelling is the only way to calculate the SSB.

4.1.2 Estimates of recruitment and stock size of fish in the North Sea and Baltic Sea

Natural mortalities of commercial fish in the North Sea and Baltic Sea are estimated by WGSAM every 3 years, and the values used as input to single species assessment. This generally changes recruitment while spawning stock is relatively unaffected as the main predation is on younger age groups. The resulting recruitment and spawning stock biomass relationships are used to estimate both B_{lim}, B_{pa} and F_{MSY} at subsequent benchmarks. The benchmarks use historic natural mortality in the estimation of F_{MSY} and hence are not dynamic.

4.2 Dynamic reference point setting

There are few cases where dynamic modelling of predators and prey is used in the estimation of reference points subsequently used in advice. This is likely due to the need to simultaneously determine targets for the stocks and thereby also deciding on trade-offs between different objectives. The group reviewed three examples, of which the first (Atlantic menhaden) is implemented in advice.

4.2.1 Accounting for predator food requirements in setting of reference levels for prey fish: Atlantic menhaden as food for and striped bass

The Atlantic States Marine Fisheries Commission sought development of Ecological Reference Points (ERP) for Atlantic menhaden *Brevoortia tyrannus* management in order to account for the role that menhaden play as a forage species. Atlantic menhaden support the largest fishery by volume on the Atlantic Coast of the United States, where harvest comprises commercial reduction and commercial bait fisheries. Menhaden serve as an important prey species for recreationally important predator species such as striped bass, weakfish, bluefish, and spiny dogfish, as well as for birds and marine mammals. The Ecological Reference Point Working Group (ERPWG) and the Menhaden Technical Committee (TC) were charged with developing ERPs. They began by developing five models of varying complexity, in addition to the single-species stock assessment model. The models included a surplus-production model with time varying r (Nesslage and Wilberg, 2019), a Steele-Henderson surplus production model (Uphoff and Sharov, 2018), a multi-species statistical catch-at-age model (McNamee 2018), an Ecopath with Ecosim model with a limited predator and prey field (Chagaris *et al.*, in review), and an Ecopath with Ecosim full model (Buchheister *et al.*, 2017a, 2017b). Based on the evaluation of trade-offs and the specified management objectives, the ERPWG put forward the single species stock assessment and the Ecopath with Ecosim model with a limited predator and prey field (NWACS-MICE) as the tools to develop ERPs for Atlantic menhaden (SEDAR 2020a; SEDAR 2020b). The ERPs were specified by the Management Board to be:

- ERP target: maximum F on menhaden that sustains striped bass at their B target when striped bass are fished at their F target, and
- ERP threshold: maximum F on menhaden that keeps striped bass at their B threshold when striped bass are fished at their F target.

The method proposed and adopted uses the Atlantic menhaden single-species assessment and the NWACS-MICE model in combination to provide reference points for menhaden management. The NWACS-MICE model was run to provide long-term, equilibrium values of F that provide for the specified ERP target and threshold. Those values of F were then used in the projections from the single-species assessment to provide a coastwide total allowable catch. Using this process, the Management Board can account for predator needs in its management and when setting the coastwide total allowable catch.

4.2.2 Herring harvest rule accounting for its role as prey

The New England Fishery Management Council (NEFMC) used MSE to develop a harvest control rule (HCR) for Atlantic herring that considered herring's role as forage (Feeney *et al.*, 2018, Deroba *et al.*, 2018). This HCR needed to balance fishing benefits and ecological services, including support of both fished and protected predators. The herring fishery also provides bait to the economically and culturally critical Maine lobster fishery. This was the first MSE conducted using a fully open stakeholder process for a US Council-managed fishery. Stakeholders identified objectives related to herring's role as forage, and identified uncertainties including environmental effects on herring productivity and growth, predator response to herring abundance, and assessment uncertainty. Management timelines constrained analytical effort to <1 year. The design was therefore simple and modular, with model components developed in parallel for herring, predators, and economics.

Eight herring operating models bracketed observed combinations of low and high herring productivity (natural mortality, M, combined with steepness, h), growth (weight at age observed in 1976–85 vs. 2005–2014), and assessment bias (0 or 60% overestimate). Implementation error and assessment error applied to the population without a formal stock assessment. Reference points were assumed known without error for each operating model-HCR combination. Time series of herring numbers and weight at age resulting from the application of each control rule were used as inputs to a simple economic model and simple deterministic delay-difference models for 3 representative predators selected by stakeholders: Bluefin tuna, common tern, and dogfish (representing groundfish). This one way flow of information could evaluate the impact of changes in the herring HCR on predators via the changed herring population indices, the main objective of an HCR accounting for herring's role as forage. However, predator models did not feed back on herring populations through mortality using this framework. Performance metrics for the herring fishery (catch and revenue stability), the herring population, and predator populations were evaluated.

Initial results were used to narrow the range of HCR types by removing three with poor performance across herring fishery, herring, and predator objectives; this occurred by consensus at a stakeholder meeting. These were constant catch, conditional constant catch, and biomass-based control rules constraining interannual changes in catch to <15%. The remaining biomass-based control rules can be selected for performance against certain metrics. For example, control rules producing consistently high tern productivity, herring biomass >90% of the MSY target, herring yield >90% of MSY, and fishery closures <1% of the time existed under all operating model assumptions (albeit with more existing in the higher productivity scenarios than in the lower productivity scenarios). This demonstrates that a range of options exists for managers to achieve objectives for herring, the herring fishery, and at least one relatively sensitive predator in this ecosystem.

While this MSE did provide information on HCRs for achieving objectives for herring, herring fisheries, and some predators, and NEFMC did select a control rule using information from this analysis, considerable improvements could be made to further explore the effects of environment, species interactions, and density dependence. Reviewers recommended testing for robustness of the HCR results to time varying herring population parameters, mis-specification of reference points, and density dependence. Further refining the species interactions by including feedbacks from predation mortality on herring, more potential relationships between herring and predators, more realistic predator population dynamics, and alternative prey was also recommended. Finally, an integrated framework considering herring supply and demand as well as direct and indirect users of herring in the ocean (predators and humans) would require considerably more data collection and investment, but would provide managers with clearer information on benefits and trade-offs associated with HCR options.

4.2.3 Multispecies MSE in the Flemish Cap area

The dynamics of the commercial species cod, redfish and shrimp in the Flemish Cap fishing ground was modelled within a multispecies gadget model, GadCap (Pérez-Rodríguez *et al*., 2017). A main goal of the modelling was contributing to the development of the NAFO roadmap for an EAF exploring alternatives to incorporate the multispecies approach into the fisheries advice process. Specifically, the multispecies model GadCap was used for the development of a multispecies MSE framework (msMSE) integrating GadCap as operating model within an a4a-MSE framework (Jardim *et al*., 2017). GadCap provides information about the "real" stocks, survey and commercial fleets that, once modified by the observation error model, are used for stock assessment in the management procedure module. Within the framework each of the three stocks has its own independent management procedure module. The current settings allow for a shortcut assessment, with and without assessment error, but also an assessment using an a4a SCAA model, that can also consider errors in the observation of survey and commercial information (Figure 7).

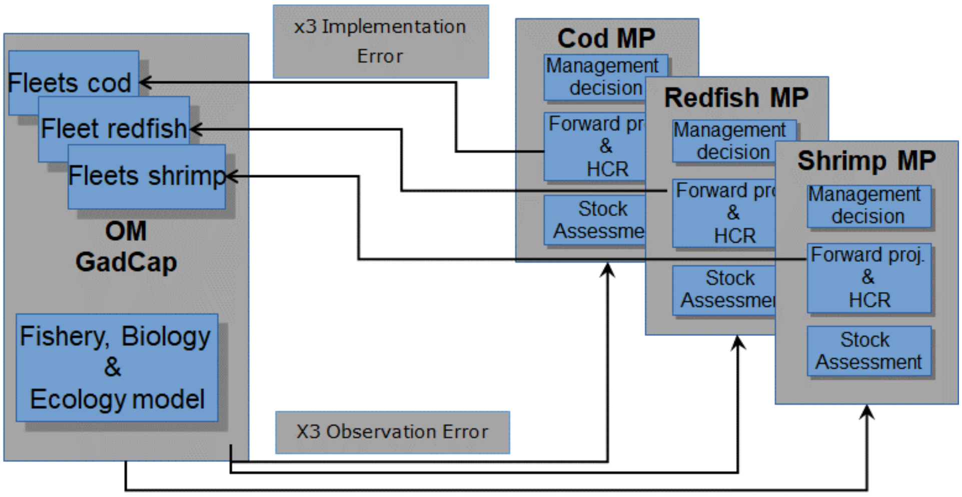

Figure 7. Multispecies gadget-a4a-MSE framework. The multispecies model GadCap was used as OM. Uncertainty on the functioning of the system was expressed by means of a stochastic SSB-Recruitment relationship in the OM. Uncertainty in the MP was simulated by introducing error in the shortcut stock assessment.

This msMSE framework was used to design and test different HCRs, where the precautionary reference points (B_{lim} and $B_{trigger}$) were estimated following the NAFO standard protocols for single species approach. The F_{target} in those HCRs were defined considering the interdependent productivity of the three stocks, i.e. from a multispecies approach. Long-term simulations were run considering multiple combinations of Fs for cod, redfish and shrimp. The results show the influence that variable fishing strategies on predators (cod and redfish) would have on the prey stocks (shrimp and redfish). Especially evident is the impact that different fishing strategies on cod would have in the productivity of shrimp and redfish. In the case of shrimp, only when very high or very low fishing pressure on cod (with resulting low stock level of redfish) is implemented, does the shrimp SSB reach values above B_{lim} (Figure 8). This pattern is due to the importance of cod as predator of redfish and shrimp, and the relevance of redfish as predator of shrimp.

The risk assessment (considering recruitment uncertainty and observation-assessment error) of the different HCRs (one stage hockey stick) showed that due to the strong trophic interactions between cod, redfish and shrimp, if shrimp is to be maintained above B_{lim}, fishing pressure on cod and redfish has to be so high that when the recruitment uncertainty is considered, the risk of being bellow B_{lim} for cod and redfish stock is very high. In conclusion, no combination of multispecies HCRs would maintain the SSB of the three stocks above B_{lim} at the same time. This result indicates that multispecies HCRs have to be designed disregarding one or two of the other species in the system. For example, if one accepts that the shrimp stock may go below the B_{lim} reference points, a number of combinations of Fs (HCRs) is obtained for which the risk of being below B_{lim} at the same time for cod and redfish below the 10% accepted by the NAFO precautionary approach framework. Additionally, as an exploratory exercise, a two-stage hockey stick HCR for cod was simulated, with the intention of testing if reducing an excessive predation capacity from high stock sizes of cod would decrease the risk of collapse of the other stocks. These two-stage HCRs were designed with an increase in fishing pressure on cod to F = 0.55 when the SSB was above 45 000 t. This two-stage HCR clearly reduced the risk of being below B_{lim} both for cod and redfish.

This study concluded that:

- It is not possible to have all three species above B_{lim}.
- Disregarding one stock (shrimp or another stock) may allow finding precautionary multispecies reference points for the others.
- The two-stage HCRs for cod reduces predation and increases probability of redfish and somehow shrimp being above B_{lim}.

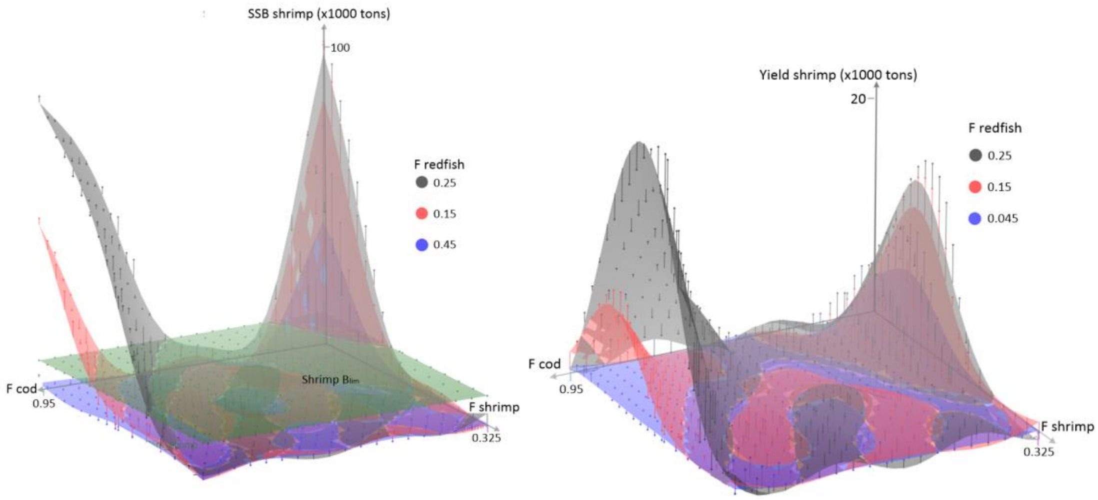

Figure 8. Estimated shrimp SSB (left panel) and yield (right panel) using the updated Flemish Cap multispecies gadget model GadCap shrimp when different F values are applied on cod (Z axis) and shrimp (X axis) for three different Fishing pressure levels on redfish.

4.3 Conclusions on predation impacts on reference points

In the ICES area, the most common inclusion of variable predation impacts in reference point estimation is the use of historic observed or modelled estimates of natural mortality in the estimation of reference points. This approach is used for Barents Sea capelin, North Sea sandeel, sprat, herring, cod, haddock and whiting stocks and Baltic Sea herring and sprat. Dynamic models are not generally used in the setting of ICES reference points, with the exception of NEA cod. The rationale for this has been that most predator stocks develop relatively slowly and hence natural mortality is unlikely to change greatly within the time span where the reference points are to be used (around 5 years). Further, the approach avoids propagating retrospective bias from the predator stock assessments into prey stock assessments by means of natural mortality changes.

The dynamic models demonstrate that in a dynamic system, the reference points are mutually dependent (Collie and Gislason, 2001). In systems with sequential exploitation of different species (e.g. a period of overfishing of predator stocks followed by a period of fishing on prey stocks) such as the Flemish Cap and the North Sea, the dynamic models generally predict that all stocks cannot be above biomass limits at the same time in the absence of considerable changes to the fishing pressure on predator stocks (see above and Kempf *et al*., 2016). Further, there are trade-offs to be made between fisheries yield of predator and prey stocks (above, Collie and Gislason, 2001, and Kempf *et al*., 2016). The need to perform these trade-offs is sometimes replaced by a sequential approach such as fixing all fishing at single-species F_{MSY} and then adapting prey species in subsequent adjustment rounds to account for the added natural mortality when having large predator stocks, effectively giving priority to predator stocks. This is equivalent to the adaptive approach taken when using static estimates of F_{MSY} and then adapting these at benchmarks as changes are observed. However, this approach does not address the need identified in the dynamic models to retain fishing pressure on predatory stocks above a minimum to retain viable stocks of prey fish. This led Kempf *et al*. (2016) to suggest a stepwise approach whereby the combinations of fishing mortality that maintain all species within safe limits are first identified and F_{MSY} then estimated among these combinations. This is based on the overarching objective to maintain all stocks at levels compatible with producing MSY by avoiding recruitment overfishing.

5 Density dependence (ToR (b)iii: stock density induced changes in stock productivity)

5.1 Is there general evidence of density-dependence in productivity?

Adequate prediction of the strength and shape of density dependence in productivity can be key to providing the best possible estimates of reference levels. Having determined unbiased estimators of the relationship between weight at age, recruitment and the most appropriate density measure, an analysis of 75 stocks from the ICES area revealed consistent density dependence in recruitment and growth of older ages. There was a tendency towards density dependence in growth of the youngest observed age group as well, but the relationship was not as strong as for older age groups. Different ecotypes of fish (pelagic, demersal and benthic) demonstrated different degrees of density dependence. Pelagic fish exhibited increasing recruitment well beyond 20% of the maximum observed SSB but showed very limited overcompensation (defined as decrease in recruitment at high SSB). In contrast, benthic and demersal species on average obtained more than 80% of their maximum recruitment at just 20% of their maximum SSB and showed greater tendency towards overcompensation. All three ecotypes showed negative correlations between weight at age of older fish and density. While the patterns were reasonably similar between ecotypes when all data were included, the strength and direction of density dependence within stocks was highly variable between time periods, with demersal stocks in particular often switching between negative and positive effects on weight at age over the time series. Hence, if we assume that B_{MSY} occurs at half virgin biomass, demersal fish may in general not exhibit decreased recruitment at biomass levels above $0.5B_{MSY}$. However, pelagic stocks are likely to experience decreased recruitment well above $0.5B_{MSY}$. A similar conclusion was reached in WKMSYREF3 (2014), where $0.5B_{MSY}$ was found to be below MSY $B_{trigger}$ in 10 out of the 19 examined stocks.

5.2 Density dependence in production models and the resulting estimates of F_{MSY}

The observed annual production against exploitable biomass for the 48 data rich ICES stocks (normalized to MSY and k, respectively, from the "general Thorson *et al.*, 2012 model") are shown in Figure 9 (Sparholt *et al.*, 2020). Here, the large variability of the production in different years is obvious, but it is also obvious that there is a clear dome-shaped relationship between surplus production and stock size, which is consistent with the classic surplus production model curves. This indicates that SPMs are reflecting observed fish population dynamics. This suggests that B_{MSY} is about 50% of B0 (the carrying capacity of the ecosystem for each particular stock). Some of the past MSEs for North Sea cod predicted very high B_{MSY} values, probably because DD in growth and natural mortality was not included in the calculations. We should expect to operate close to the observed span of SSB values when deciding on appropriate SSB reference points.

The results highlight the errors which can be introduced by neglecting density dependence in the calculation of F_{MSY}, and suggest that surplus production models are one potential modelling tool to remedy this.

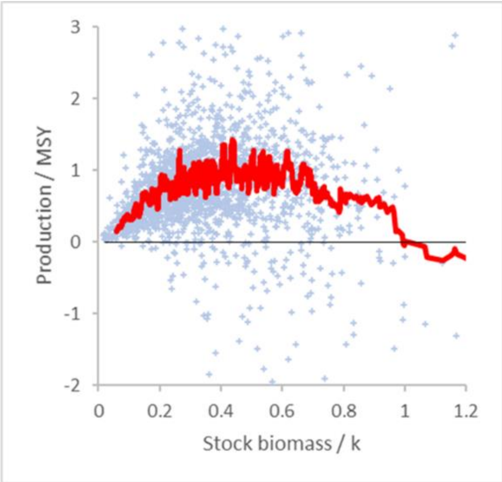

Figure 9. Surplus production vs. stock biomass, normalized to MSY and *k* (carrying capacity), respectively, for 48 data-rich stocks. For clarity, 34 out of 1901 data-pairs were not included because they were outside the intervals on the y-axis but were quite evenly spread around the general pattern. The red line is a running mean of 25 points. The "general Thorson *et al*. (2012)" model (φ = 1.736) was used to estimate MSY and *k* by stock.

5.3 Approaches to adapting to density dependent changes between benchmarks

The effect of density-dependent growth is most important at higher stock sizes. To account for the decrease in productivity at large stock size of Northeast arctic cod, F_{target} depends on stock size. This is done by including a double hockey stick in the HCR, as illustrated in fig. 10. The change from a flat F above B_{pa} to the present HCR was made in 2016. The double hockey stick allows higher fishing pressure at high stock sizes where density dependence decreases productivity, but keeps F lower at lower stock sizes, thus there is no increased risk of stock collapse. The change points and F levels in the HCR are based on expert judgement, but can also be evaluated using quantitative methods.

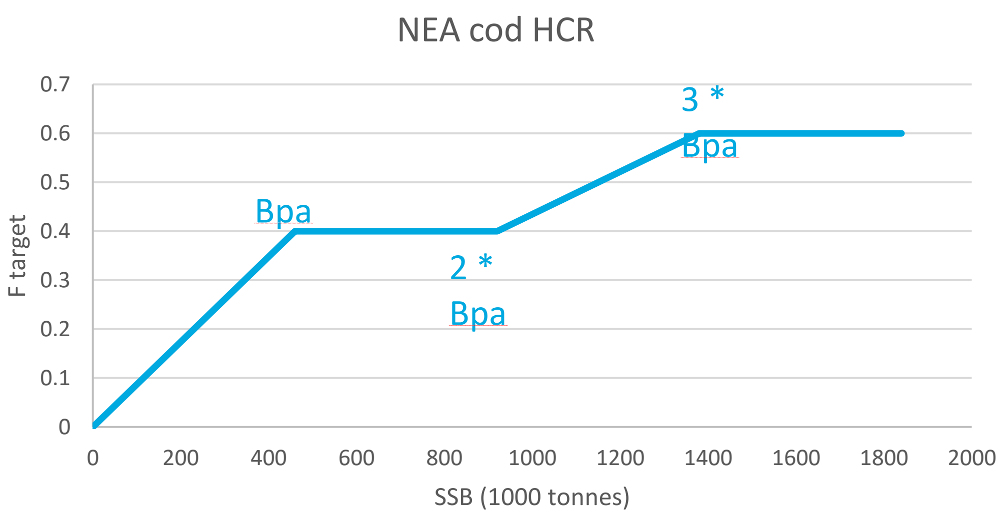

Figure 10. Relationship between stock biomass and F_{target} of Northeast Arctic cod. The target F values are based on F_{MSY} with cannibalism (F = 0.4) and with the addition of density-dependent growth (F = 0.6).

5.4 Conclusions on density-dependent effects on reference points

WKMSE2 noted in the guidelines for MSE evaluations that "A critical part of designing any MSE exercise is to identify early on, which key processes need to be included in the operating model(s)" (ICES, 2019c). Hence, where density dependence is an important driver of stock dynamics at current biomass levels (see below), it should be included in the evaluation of reference points, either through a full HCR evaluation or through the simpler ICES procedure using EqSim.

Density dependence in the recruitment function are already included in reference point estimations, and are key to the estimation of B_{lim}. This practice should continue. However, as seen in the summaries above, density dependence may also be an important process later in the life of the fish and therefore impact on estimation of F_{MSY} and other F reference points. Density dependence is used as an approximation to model food limitation and occasionally cannibalism in stocks. In general, food limitation can be influenced by variations in food availability, in competition, and in biomass of the stock under consideration. Only the last of these is accounted for by density-dependent growth. It is important to bear this limitation in mind when including density dependence within a simulation model, and evaluate the degree to which density dependence is likely to be able model food limitation. Where the variation in change in food per predator is driven primarily by fluctuations of the biomass of the stock under evaluation, then density dependence is likely to be a good approximation to this process. Where there are significant variations in the prey biomass, or major and varying competition pressure for food, then density dependence may not be able to appropriately approximate food limitation, in which case prey abundance and competition between species will need to be modelled in more detail (see ToR b ii). Examples of where density dependence in a single species is unlikely to be a viable model include competition for food between herring and sprat in the Irish Sea (e.g. Bentley *et al.*, 2020); between NSS herring, blue whiting and mackerel in the North Atlantic (Bachiller *et al.*, 2016); or competition for krill between minke and humpback whales in the South Atlantic (Konishi and Walloe, 2015).

Where density dependence is evident in a stock, it can have a large impact on the estimated reference points. An analysis for NEA cod using the PROST software, which allows for including density dependence (Sparholt *et al.*, 2020) indicated $F_{MSY} = 0.2$ where density dependence was only in recruitment, $F_{MSY} = 0.4$ when including density dependent cannibalism, and $F_{MSY} = 0.6$ when also incorporating density dependent growth.

There are two possible approaches to deal with density dependence in management. One solution to this is to evaluate F_{MSY} for a range of stock sizes including density dependence if it is considered important over the likely lifespan of the reference points. This is likely to be robust but will provide you with high levels of F_{MSY} at both low and high biomass. These high F target levels can be relevant for stocks which are expected to experience density-dependent effects in the period which the reference points are to be applied for. They will, however, present an increased risk for stocks which are at low levels and are expected to rebuild slowly. A second approach is to apply a double hockey stick HCR. A double hockey stick HCR applies a second increase in target F at high stock size. This ensures that the higher F_{target}, accounting for density dependence, only applies at high stock sizes. This double hockey stock approach is more precautionary than having a constant higher F_{target} at low stock size and is robust to productivity changes at high stock size. It may thus be more appropriate in cases where the stock size is expected to vary significantly over the lifespan of the reference point. The double hockey stick acts to push stock size towards a given biomass range, which may also be useful in multispecies management.

6 A stepwise approach to making appropriate changes in target, range, and limit reference points (ToR c)

The discussion on reference points focussed on four topics: (1) major sources of uncertainty in reference points, (2) the time period for which reference points should reflect MSY objectives and precautionary risks, (3) how to obtain the best predictions of productivity for this period, (4) and how to use different predictive models to estimate reference points and harvest control rules. These discussions are summarized with guidelines about when and how to change reference points.

6.1 Major sources of uncertainty in reference points

Biological reference points depend on life-history processes—mortality, growth, maturity, recruitment—and the demographic parameters that are used to quantify these processes. Although ICES has clear guidelines for the estimation of reference points (Section 2.1), uncertainty remains regarding which processes to include, the estimation of demographic parameters, and whether these parameters change over time.

Recruitment remains the most variable component of productivity, varying inter-annually by several orders of magnitude in most stocks. As a result, it is the aspect most difficult to predict adequately. One of the difficulties arising is that parametric simulation methods of all kinds will predict occasional years with unprecedented high recruitment unless upper limits are specified directly. This problem persists regardless of the type of model (moving average, stock-recruitment relationship with error, etc.). In addition, recruitment failures combined with high fishing pressure have historically led stocks to decline to levels that are not always followed by stock recovery even when fishing is subsequently reduced. High fishing pressure and non-fishing impacts on recruitment can be confounded and it can be hard to separate the effects of one from the other. These considerations make stock-recruitment based reference points both highly uncertain and highly influential on sustainability.

Most MSY estimation methods presented to the group included a stock-recruitment relationship that was non-proportional and hence density dependent. However, some cases investigated recruitment at different levels (alternate regimes) with no specification of the effect of stock size on recruitment. It should be stressed that such models will not reflect the excepted decline in recruitment at low stock size and hence may underestimate the need for precautionary limits to fishing mortality.

The “hockey-stick” model provides reliable estimates of B_{lim} for stocks with low CV in recruitment (Section 2.4). For stocks with high CV or small contrast in stock size, the estimation of B_{lim} and MSY based reference points is a major source of uncertainty. While the discussion of uncertainty in the stock-recruitment relationship is mostly centred on limit reference points (B_{lim}), uncertainty in the stock-recruitment relationship is equally influential in the estimation of MSY related reference points. Hence, the recommendation to shorten recruitment time series only when presented with substantial evidence (see below) is equally relevant for MSY based reference points, including all levels of minimum stock size based on B_{MSY} or virgin biomass. The US limit reference point for biomass (0.5*B_{MSY}) is not directly related to the ICES reference point B_{lim}. It would appear from the analysis of density dependence that the 0.5*B_{MSY} threshold is unlikely to safeguard recruitment of a large proportion of the stocks (Section 5.1).

6.2 Determining the period for which reference points should reflect precautionary, MSY and ecosystem objectives

In general, ICES re-evaluates precautionary and MSY reference points in benchmarks every 5 years, though with the pressure on the ICES system, this interval is sometimes extended up to 10 years. Hence, any reference points estimated should reflect the conditions expected in the next 5–10 years. Long-term simulations are often used to estimate single-species reference points such as F_{MSY}, which are then applied in the medium term. In an ecosystem context, reference points are conditional on current environmental conditions, including productivity and the abundance of interacting species. Even the level of cannibalism, which can be considered a single-species process, depends on levels of alternative prey species (Gislason, 1999; Collie and Gislason, 2001). Therefore, care should be taken to ensure that reference points reflect changes that are likely to occur within a 10-year time period, conditioned on management decisions for interacting stocks.

A related issue is choosing the time window of data for calculating reference points, which is critical for projecting future recruitment levels. While using the entire time series to fit a stock-recruitment model should provide more precise parameter estimates, a more recent time window may better reflect current environmental conditions. This dilemma can be contentious when it affects the determination of stock status and recommended harvest levels.

How do we distinguish "exceptional" conditions from what is "normal"? The WK did not provide a specific answer to this question, but there was consensus that strong evidence of change is required before taking action to change reference levels. Unless the variation around the stock recruitment relationship is low, substantial evidence should be available before truncating the time series in the estimation of stock-recruitment based reference points. Preferably, causal understanding or statistical tests of whether a change in inflection point, recruitment level or both has occurred should be used to guide the decision. The entire time series is needed to demonstrate whether a regime shift occurred (e.g. different levels of recruitment at the same level of SSB, Walters 1987).

Separating signal from noise is important in choosing the time window for calculating reference points. Changes in the prevailing environmental conditions are difficult to discern from stock data alone, even with good data. Looking across multiple stocks or at auxiliary sources of environmental information may provide better indicators when environmental conditions have changed for a particular stock (Tableau *et al.*, 2019).

6.3 Obtaining the best predictions of productivity

There are several components of productivity. It is therefore important to be clear about which part of the life cycle is believed to be changing, and to tailor any response to this knowledge. For example, in the recent benchmark of Eastern Baltic cod (ICES, 2019d), productivity was described as the combination of four different process: recruitment, growth, condition and natural mortality and CUSUM methodology was applied to time series of the key parameters describing the four processes in order to identify possible shifts in the time series. For NSAS herring, time series of R/SSB were used.

Variation in weight at age, maturity and natural mortality tends to be positively autocorrelated in most stocks, presumably reflecting gradual changes in feeding conditions and predator densities. In contrast, recruitment success shows positive or negative autocorrelation depending on the stock. Weight at age and maturity can be measured each year; as long as they develop slowly, recent averages can be used to update reference points. However, if density dependence is demonstrated in growth or maturity, these processes may need to be incorporated in the reference point calculation. The natural mortality rate is not measured directly but needs to be estimated from life-history proxies or predator-prey models. Since the natural mortality rate generally changes slowly as predator populations develop, it can in many cases be projected for 5 to 10-year periods and updated as necessary at benchmark assessments.

Models of trophic processes need to be based on empirical evidence of the interaction. Evidence of predation (variable *M*) is much more common and widespread than bottom-up food limitation, which explains why multispecies models focus on predation. This choice is partly because predation is more directly observable than food limitation, but also because predators in many temperate ecosystems have many potential prey species. Shifts in trophic interactions can result from changes in species abundance or changes in their distribution. To date, multispecies models have been more successful in accounting for variations in abundance than variations in distribution. The WK encourages more ICES work to improve spatial modelling and species overlap.

In contrast to the progress made on incorporating variable predation mortality in stock assessments and reference points, food limitation has been more difficult to incorporate. Where the food limitation is due to the biomass of the single predator species, it can be handled through density-dependent growth (Section 5). However, food limitation can also arise due to depletion of the prey species, or through competition resulting in food limitation for one or more of the competing species. Examples include Celtic Sea herring experiencing competition for food; Irish Sea herring and sprat competing for food; minke and humpback whales competing for food in the Southern Ocean (Konishi and Walløe, 2015), and NSS herring, blue whiting, and mackerel which compete for similar food and are being investigated in the Norwegian SIS Harvest project.

There is strong evidence for density-dependence in recruitment and for density-dependent growth post recruitment for some stocks (Section 5.1). Maturation may also be affected by food limitation but the data to test this are sparse. Where it occurs, density-dependent growth can be very important for the F_{MSY} estimate (e.g. NEA cod). Density-dependent growth is a proxy for food limitation, which in reality depends on variations in food and competition between species as well as variations in predator biomass. Weight at age may change over time for reasons other than food limitation (e.g. temperature, fisheries selectivity). Evidence of food limitation is therefore needed to support estimates of density-dependent growth.

The best prediction of productivity depends on the understanding of the shape and direction of relationships (e.g. density, temperature and food availability) and the predictive ability of the relationship. A choice must be made between predicting a distribution of future conditions with functional relationships between covariates and productivity or by statistical methods such as time-series models or other types of relationships between the historic data and the future predictions without including independent information. This choice should be made based on a combination of understanding of the ecosystem and predictive ability.

6.4 Estimating reference points

Include all relevant and important processes

An early stage of reference point calculations should be a scoping exercise to identify which processes and variabilities to include in the estimation. A list of potential processes could be provided. Identifying a limited number of key drivers enables the important processes to be included and streamlines the development of the estimation.

In general, management aims at reaching all target and limit reference points for all stocks simultaneously. It is frequently not possible to attain MSY $B_{trigger}$ and/or B_{MSY} for all species simultaneously, and moreover, in situations where there are strong species interactions, even attaining biomass limit reference points for all species may not be possible. The Flemish Cap example described in 4.2.3 shows one case where B_{lim} cannot be maintained for all commercial species simultaneously. An overall understanding of the processes is therefore important for identifying overall management objectives as well as estimating individual reference points.

Where ecosystem processes can be directly incorporated into the tactical models (mostly single-species) for calculating reference points, this is the preferred approach. One example of this would be to include density dependence in e.g. growth directly into the population models used in an MSE or into EqSim. Where the ecosystem variability cannot be directly included in the single-species models, ICES can use strategic ecosystem models to incorporate modest adjustments to target reference points, while ensuring precautionarity (and if required good yield criteria). For example, the Atlantic menhaden F has been conditioned on striped bass abundance (Section 4.2.1) and reference points in the Irish Sea conditioned on ecosystem drivers (Section 3.1). This approach is feasible as long as the adjustments to the target F remain within the F_{MSY} range. In general, the ability to "tune" F_{target} within existing limit and potentially range reference points gives the flexibility to include some ecosystem variation into the quota advice without requiring re-estimation of the reference points (Section 3.3). In other situations, ecosystem tuning implies F_{target}s that fall outside the single-species precautionary ranges (e.g. in the Flemish Cap example it wasn't possible to achieve B_{lim} for all species simultaneously).

Errors resulting from truncating time series

Analysts need to understand uncertainty surrounding the productivity of a stock and the risk/rewards associated with updating reference points to match potential changes in productivity. This issue is mainly aimed at methods whereby the window for estimating productivity changes (e.g., stock-recruitment relationship based on most recent 10 years vs full historical range of data). The choice of recruitment window depends on 1) the length of the recruitment time series, 2) the quality of recruitment estimates, 3) the CV in recruitment, 4) the time scale over which hypothesized (or empirical) productivity changes occur (which itself may be related to oceanography, climate, species generation time, predator/prey dynamics, etc.), and 5) how the stock responds to these changes. The magnitudes of these uncertainties determine whether choosing a more recent recruitment window is useful for changing reference points or whether uncertainty will overwhelm any expected utility, creating situation in which the management system is chasing its own tail (due to time lags) or advice falls into a "shifting baseline syndrome" sort of trap.

Is it a good idea to change to a lower B_{lim} in a low productivity regime or should a higher recovery target be maintained? The answer depends on whether a persistent change is considered to have happened. There are two types of error associated with choosing a recruitment time window. If the null hypothesis is a stationary time series (no shift in productivity), changing the recruitment window and associated reference points would be Type-I error. Conversely, failing to change the reference points when productivity shifts corresponds to a Type-II error. Participants were mostly concerned about Type-I errors, especially for declining and depleted stocks. There is a risk that lowering the reference points could trap a stock at low abundance and prevent rebuilding. For increasing stocks there is less risk of Type-I errors because most harvest control rules hold F constant at high stock size. There are examples of stocks with measured increases in productivity but there is less concern about foregone catches if the reference levels are not changed in response (Type—II error). This asymmetry occurs because the fishery management system is inherently risk-averse-stocks can be overfished rapidly but take years to rebuild. In some jurisdictions, laws require rebuilding plans for overfished stocks but do not mandate reducing stocks to achieve MSY.

Situations should be avoided whereby the determination of a lower productivity regime would result in a higher Total Allowable Catch (TAC). This situation could arise in the US if stock status changes so that it is no longer considered overfished and rebuilding requirements are thus relaxed. The exact consequences of making Type-I or Type-II errors depend on the harvest control rule that is used to set the TAC and especially the action taken at low stock size (Figure 1). If fishing mortality is reduced at low stock biomass, a depleted stock may still be able to rebuild, whether or not productivity and the corresponding reference points have changed. Changing the reference points will change the biomass thresholds and fishing mortality targets but not the underlying shape of the harvest control rule. Even a generic HCR may be robust to changing productivity (Kritzer *et al*., 2019).

The role of simulation

Stochastic simulations are typically used to estimate F_{MSY} and, in the US, the corresponding B_{MSY}. Care should be taken to include all relevant processes in the simulation with stochastic errors that match the observed patterns. When running simulations, care is needed to ensure that the processes in the simulations remain within the range of the tuning data (both the observed range of values, and the autocorrelation structure). This is an issue with simulation of recruitment in current programs. For many stocks, standard programs (e.g. EqSim) will suffice. For stocks subject to environmental change, additional processes may need to be incorporated in customized simulations. The impacts of correlations between different productivity components should be included whenever these prove to be substantial, as these may alter precautionary reference points substantially.

Simulations can also be run to understand the risks/rewards associated with decisions to truncate time and update reference points (B_{lim}, etc.). Haltuch and Punt (2011) quantified the occurrence of Type-I (identifying a change in productivity when there is not one) and Type-II errors (failing to identify environmental impact on productivity when there is one) for multiple life stages. Based on their simulations, these errors can be common. This type of analysis could inform situations when “default” reference points might be appropriate and when dynamic ones need to be developed.

Management strategy evaluation (MSE) style simulations can be used to evaluate the performance of harvest rules for stocks subject to environmental change, but this level of analysis is not required for every stock. Complex MSEs should be reserved for ecosystem-level questions with poorly specified objectives, complex interactions, and trade-offs between user groups. Since many of the results are transferable, analysts should take advantage of lessons learned from prior studies on other stocks.

It is better to link existing software tools than to re-invent the wheel. This approach will reduce the development time, reduce the likelihood of software errors, and make it easier to keep the different sections up to date. Work at this meeting (Section 4.2.3) presented an example of this approach through linking a multispecies Gadget model to the A4A MSE framework to evaluate multispecies HCRs.

6.5 Stepwise guidelines on changing reference points

A. When should reference points be re-evaluated?

Biological reference points should be re-estimated at benchmark assessments, which occur on a roughly 5-year cycle. This timescale matches the management system, avoids “whipsaw” changes 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.

B. What are the signals of change and how can they be tested?

The benchmark assessment provides an opportunity to look for changes in the population dynamics beyond those already incorporated in the stock-assessment model. Such changes may occur in the mean or variance of per-capita recruitment, weight at age, natural mortality, or maturity. These changes could be indicated by bias in the assessment or forecast, or patterns in residuals. Biomass not responding to changes in fishing pressure or recruitment not responding to changes in biomass suggest additional drivers not included in a single-species assessment model. Multiple species in an ecosystem experiencing concurrent low recruitment success may be caused by strong ecosystem change. Other ecosystem signals of regime change include a step change in zooplankton or other trophic groups, shifts in temperature or other physical habitat metrics.

Statistical methods exist to test for significant changes from the null hypothesis of stationarity. These include likelihood ratio tests for dynamic stock-recruitment models and regime shift indicators. If there are no significant changes (the variability is white noise or high-frequency autocorrelation without trend) standard procedures can be used to update the reference points. Otherwise change and re-evaluate reference points as described in C.

C. How should reference points be changed and evaluated?

Is there a mechanistic process linking productivity and explanatory variables?

Recruitment: Is the observed change in the density-independent (e.g. slope at the origin) or density-dependent parameter (e.g. maximum recruitment)? For example, if all recruitment habitat can support half the previous recruitment, the shape of the S-R relationship and hence B_{lim} remains the same but recruitment at any SSB is halved. If the extent of recruitment habitat is halved, the SSB required to fill this habitat would only be half the previous B_{lim}. The workshop considered that more evidence should be required to change precautionary biomass reference levels derived from stock recruitment relationships than to change F_{MSY}. This distinction would be automatic if we use the entire S-R time series to estimate B_{lim} and moving windows for F_{MSY}.

Growth and maturity: Weight at age and sometimes maturity are measured annually and hence can easily be adjusted in reference point calculations. It is important to test if there are statistical relationships between density and these factors in historic data. If so, density dependence can be incorporated in the reference point algorithm.

Cannibalism: For some species, density dependence occurs through cannibalism, in which case it may be possible to use a multispecies model or a single species model extended to include cannibalism to relate the density of larger fish to mortality of smaller fish. In this case, the mortality is determined by the fishing mortality and the abundances of alternative prey species. In other cases, fishing mortality on the species in question does not regulate natural mortality.

Mortality: Changes in natural mortality (*M*) can affect reference points, especially if it co-varies with growth. Natural mortality can be forecast for a 5–10 year-period assuming existing harvest rules and abundance of interacting species persist. However, *M* is not routinely estimated in stock assessments.

The prediction error variance can be used to determine whether using the causal explanatory variables provides better predictions than a stationary model. If so, can the causal variable be predicted in advance or does the forecast depend on future measured values of the causal variable? In the latter case, what is the time lag between measuring the causal variable and predicting fish productivity (e.g. cold pool volume affects recruitment two years later)?

If a causal link has been hypothesized as influential but not established, more research may be warranted even if reference points can be adjusted on a purely empirical basis. Causal understanding is especially important if future data will fall outside the envelope of historical data (e.g. climate change). A mechanistic understanding also provides decision makers assurance that they are adjusting reference points in the right direction.

When causal mechanisms for change are unknown, the data series can be examined with statistical tools (e.g. time series or state-space models) to determine if there are temporal patterns that can be used for prediction. Time series that shift once do not contain much information on the probability of shifting between regimes and the probability of shifting in one direction may differ from the probability of shifting in the other direction. Some regime shifts are actually responses to more gradual development in climatic conditions (e.g. NAO). In these cases, more frequent updating is required to ensure that more knowledge is incorporated as it is collected.

Regional benchmark assessments of multiple species (e.g. WKBALT, WKACT, etc.) may help to identify which species are responding to common environmental signals. Standardized computer software can be used across species and modified as appropriate. This approach increases the efficiency of analysis, review, and understanding by managers. Multispecies assessments show which species will "win" and "lose" under systemic environmental change. Do ecosystem-based reference points fall within the existing limits or is it impossible to meet all single-species limits simultaneously? If the latter, what trade-offs between species and fleets are acceptable?

In the ICES management system, there is a clear difference between F and precautionary biomass reference points. Failing precautionary reference points are likely to carry a large cost of recruitment failure or large uncertainty in the chance of stock recovery. This distinction does not exist in the US system where all reference points are treated as limits.

A first requirement for including density dependence to become routine is that this functionality is included in the existing EqSim program within ICES, as well as in the commonly used MSE tools within the ICES area. With this capability, F_{MSY} can be estimated with and without density-dependent growth. Where substantial density dependent effects are expected to occur and remain relatively constant over the lifespan of the reference point, then the F_{MSY} including density dependence should be applied to the ICES hockey stick HCR. Where large changes of stock size are expected to cause variations in the strength of the density dependence over the lifespan of the reference point, then a double hockey stick can be used with F_{MSY} estimated without density dependence being the F_{target} used at medium stock size and F_{MSY} estimated with density dependence the F_{target} used at high stock size. This approach will enhance precautionarity at low stock size without foregoing yield at high stock size.

D. General Guidance

Ecosystem functioning is complex, and no one set of guidelines will be appropriate to every situation. These guidelines must therefore not be used to constrain possible responses in changing reference points and providing fisheries advice. It is important that the biomass limit and the fishing pressure reference points reflect the environmental drivers for each stock where this can be justified and documented by the best available science. We note that the existing benchmark procedure gives scope to provide an independent review of any particular approach. Progress toward ecosystem-based fisheries management is a continuing process. Therefore, these guidelines cover elements that could be incorporated into the advice process now, and ones that require further research. We also stress that, although acquiring a mechanistic understanding is important, a lack of complete understanding should not delay our ability to react to sudden and persistent changes. The process for estimating and updating reference points can only be fully evaluated in the context of the harvest control rule and the advice that is based on it; this broader evaluation is the logical next step.

6.6 Specific recommendations:

1. Any estimation of reference points should include an evaluation of which processes are critical over the lifespan of the reference point, and should include density dependence in the simulations if that is assessed as important at the stock sizes likely to apply over the lifetime of the reference point.
2. One task for any benchmark is to evaluate if previous assumptions about the critical processes (potentially including density dependence) remain valid given the stock development, and to respond appropriately.
3. Density dependence in recruitment should continue to be accounted for through the B_{lim} value which the HCR aims to keep the stock biomass above. Since a bias can be introduced by shortening the time-series, B_{lim} estimates based on truncated time series or adjusted for productivity should be used with great care.
4. ICES should include the ability to have density dependence in the EqSim program for evaluating reference points.
5. Where environmental conditions or density dependent impacts are likely to change substantially over the lifespan of the reference point(s), the possibility of varying those reference points should be considered. For density dependent effects this could be accomplished through a double hockey stick HCR, thereby eliminating the necessity for inter-benchmark adjustment of F_{MSY}. Small variations based on ecosystem drivers can be incorporated into the quota advice by similar pre-defined adjustments to F_{target}.
6. Where reference points are allowed to vary with stock size or environmental conditions, these changes should be used as the basis for any determination of stock or fishing status (e.g. the ICES traffic light system)
7. The process for estimating and updating reference points needs to be evaluated in the context of the harvest control rule and the advice that is based on it, in particular where F_{MSY} is used as the F_{target} in a HCR.

7 References

Bachiller, Eneko & Skaret, Georg & Nøttestad, Leif & Slotte, Aril. (2016). Feeding Ecology of Northeast Atlantic Mackerel, Norwegian Spring-Spawning Herring and Blue Whiting in the Norwegian Sea. PLOS ONE. 11. e0149238. 10.1371/journal.pone.0149238.

Barrowman, N. J., & Myers, R. A. (2000). Still more spawner-recruitment curves: the hockey stick and its generalizations. Canadian Journal of Fisheries and Aquatic Sciences, 57(4), 665-676.

Bentley, J. W., Hines, D.E., Borrett, S.R., Serpetti, N., Hernandez-Milian, G., Fox, C., Heymans, J.J., & Reid, D.G. (2019a). Combining scientific and fishers knowledge to co-create indicators of food web structure and function. ICES Journal of Marine Science 76(7): 2218-2234.

Bentley, J. W., Serpetti, N., Fox, C., Heymans, J.J., & Reid, D.G. (2019). Fishers knowledge improves the accuracy of food web model predictions. ICES Journal of Marine Science 76(4): 897-912.

Bentley, Jacob & Serpetti, Natalia & Fox, Clive & Heymans, Johanna & Reid, Dave. (2020). Retrospective analysis of the influence of environmental drivers on commercial stocks and fishing opportunities in the Irish Sea. Fisheries Oceanography. 10.1111/fog.12486.

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

Buchheister, A., Miller, T. J., and Houde, E. D. (2017a). Evaluating ecosystem-based reference points for Atlantic menhaden. Mar. Coast. Fish. 9(1), 457-478.

Buchheister, A., Miller, T. J., Houde, E. D., and Loewensteiner, D. A. (2017b). Technical Doc-umentation of the Northwest Atlantic Continental Shelf (NWACS) Ecosystem Model. Report to the Lenfest Ocean Program, Washington, D.C. University of Maryland Center for Envi-ronmental Sciences Report TS-694-17. Available at: http://hjort.cbl.umces.edu/NWACS/TS_694_17_NWACS_Model_Documentation.pdf

Chagaris, D., Drew, K., Schueller, A., Cieri, M., Brito, J., and Buchheister, A. (In Review). Eco-logical reference points for Atlantic menhaden. Frontiers in Marine Science.

Clausen, L. W., Rindorf, A., van Deurs, M., Dickey-Collas, M., & Hintzen, N. T. (2018). Shifts in North Sea forage fish productivity and potential fisheries yield. Journal of Applied Ecology, 55(3), 1092-1101.

Collie, J. S., & Gislason, H. (2001). Biological reference points for fish stocks in a multispecies context. Canadian Journal of Fisheries and Aquatic Sciences, 58(11), 2167-2176.

De Oliveira, J. A. A. 2013. North Sea cod evaluations. ICES Document CM 2013/ACOM 72: 74. pp.

Deroba, J. J., Gaichas, S. K., Lee, M.-Y., Feeney, R. G., Boelke, D. V., and Irwin, B. J. 2018. The dream and the reality: meeting decision-making time frames while incorporating ecosystem and economic models into management strategy evaluation. Canadian Journal of Fisheries and Aquatic Sciences, 76(7), 1112-1133.

DiCosimo, J., Methot, R. D., and Ormseth, O. A. 2010. Use of annual catch limits to avoid stock depletion in the Bering Sea and Aleutian Islands management area (Northeast Pacific). ICES Journal of Marine Science, 67: 1861–1865.

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.

DFO. 2013. Proceedings of the National Workshop for Technical Expertise in Stock Assessment (TESA): Maximum Sustainable Yield (MSY) Reference Points and the Precautionary Approach when Productivity Varies; December 13-15, 2011. DFO Can. Sci. Advis. Sec. Proceed. Ser. 2012/05

Electronic Code of Federal Regulations (eCFR). https://www.ecfr.gov/ (Accessed 21 September 2020). https://www.ecfr.gov/cgi-bin/text-idx?SID=71b8c6026001cb90e4b0925328dce685&mc=true&node=se50.12.600_1310&rgn=div8

Feeney, R. G., Boelke, D. V., Deroba, J. J., Gaichas, S. K., Irwin, B. J., and Lee, M.-Y. 2018. Integrating Management Strategy Evaluation into fisheries management: advancing best practices for stakeholder inclusion based on an MSE for Northeast U.S. Atlantic herring. Canadian Journal of Fisheries and Aquatic Sciences, 76(7), 1103-1111.

Gislason, H. (1999). Single and multispecies reference points for Baltic fish stocks. ICES Journal of Marine Science, 56(5), 571-583.

Haltuch, M. A., & Punt, A. E. (2011). The promises and pitfalls of including decadal-scale climate forcing of recruitment in groundfish stock assessment. Canadian Journal of Fisheries and Aquatic Sciences, 68(5), 912-926.

Hilborn, R. (2010). Pretty good yield and exploited fishes. Marine Policy, 34(1), 193-196.

ICES. 2014. Report of the Joint ICES-MYFISH Workshop to consider the basis for F_{MSY} ranges for all stocks (WKMSYREF3).

ICES. 2015. Report of the Workshop to consider F_{MSY} ranges for stocks in ICES categories 1 and 2 in Western Waters (WKMSYREF4). ICES CM 2015/ACOM:58

ICES. 2017a. ICES fisheries management reference points for category 1 and 2 stocks. ICES Advice, Book 12, Section 12.4.3.1.

ICES. 2017b. Report of the Benchmark Workshop on the Irish Sea Ecosystem (WKIrish3), 30 Jauary–3 February 2017, Galway, Ireland. ICES CM 2017/BSG:01. 220 pp.

ICES. 2017d. Report of the Benchmark Workshop on Pelagic Stocks, 6–10 February 2017, Lisbon, Portugal. ICES CM 2017/ACOM:35. 278 pp.

ICES. 2018a. Report of the Benchmark Workshop on Pelagic Stocks (WKPELA 2018), 12–16 February 2018, ICES HQ, Copenhagen, Denmark. ICES CM 2018/ACOM:32. 313 pp.

ICES. 2018b. Report of the Workshop on stakeholder input to, and parameterization of, ecosystem and foodweb models in the Irish Sea aimed at a holistic approach to the management of the main fish stocks (WKIrish4), 23–27 October 2017, Dún Laoghaire, Ireland. ICES CM 2017/ACOM:54. 35 pp.

ICES. 2019a. Report of the Workshop on an Ecosystem-based Approach to Fishery Management for the Irish Sea (WKIrish5), 5–9 November 2018, Dublin, Ireland. ICES CM 2018/ACOM:66. 55 pp.

ICES. 2019b. Workshop on the Iberian Sardine Management and Recovery Plan (WKSARMP). ICES Scientific Reports. 1:18. 168 pp. http://doi.org/ 10.17895/ices.pub.5251

ICES. 2019c. Workshop on Guidelines for Management Strategy Evaluations (WKGMSE2). ICES Scientific Reports. 1:33. 162 pp. http://doi.org/10.17895/ices.pub.5331

ICES. 2019d. Benchmark Workshop on Baltic Cod Stocks (WKBALTCOD2). ICES Scientific Reports. 1:9. 310 pp. http://doi.org/10.17895/ices.pub.4984

ICES. 2020a. Guide to ICES advisory framework and principles. *In* Report of the ICES Advisory Committee, 2020. ICES Advice 2020, section 1.1. https://doi.org/10.17895/ices.advice.7648

ICES. 2020b. Workshop on an Ecosystem Based Approach to Fishery Management for the Irish Sea (WKIrish6; outputs from 2019 meeting). ICES Scientific Reports. 2:4. 32 pp. http://doi.org/10.17895/ices.pub.5551

Jardim, E., Scott, F., Mosqueira Sanchez, I., Citores, L., Devine, J., Fischer, S., Ibaibarriaga, L., Mannini, A., Millar, C., Miller, D., Minto, C., De Oliveira, J., Osio, G., Urtizberea, A., Vasilakopoulos, P. and Kell, L., Assessment for All initiative(a4a) - Workshop on development of MSE algorithms with R/FLR/a4a, EUR 28705 EN, Publications Office of the European Union, Luxembourg, 2017, ISBN 978-92-79-71290-6, doi:10.2760/18924, JRC106750

Kempf, A., Mumford, J., Levontin, P., Leach, A., Hoff, A., Hamon, K. G., ... & Smout, S. (2016). The MSY concept in a multi-objective fisheries environment–Lessons from the North Sea. Marine Policy, 69, 146-158.

Konishi, K. and Walløe, L. (2015). Substantial decline in energy storage and stomach fullness in Antarctic minke whales during the 1990s. Journal of Cetacean Research andManagement15, 77–92

Kritzer, J.P., Costello, C., Mangin, T., Smith, S.L. 2019. Responsive harvest control rules provide inherent resilience to adverse effects of climate change and scientific uncertainty. ICES. J. Mar. Sci. 76(6):1424-1435.

McNamee, J.E. (2018). A multispecies statistical catch-at-age model for a Mid-Atlantic species complex. [dissertation]. [Narragansett (RI)]: University of Rhode Island Graduate School of Oceanography. Available at: https://digitalcommons.uri.edu/oa_diss/758/

Methot, R. D., Tromble, G. R., Lambert, D. M., and Greene, K. E. 2014. Implementing a science-based system for preventing overfishing and guiding sustainable fisheries in the United States. ICES Journal of Marine Science, 71: 183–194.

Miller, T.J., O'Brien, L, and Fratantoni, P.S. 2018. Temporal and environmental variation in growth and maturity and effects on management reference points of Georges Bank Atlantic cod. Can. J. Fish. Aquat. Sci. 75: 2159-2171.

Minto, C., Mills Flemming, J., Britten, G.L., and Worm, B. 2013. Productivity dynamics of Atlantic cod. Can. J. Fish. Aquat. Sci. 71(2): 203–216. doi:10.1139/cjfas-2013-0161.

Nesslage, G.M. and Wilberg, M.J. (2019). A performance evaluation of surplus production models with time-varying intrinsic growth in dynamic ecosystems. Can. J. Fish. Aquat. Sci. 76(12), 2245-2255 DOI: 10.1139/cjfas-2018-0292.

Pérez-Rodríguez, A., D. Howell, M. Casas, F. Saborido-Rey, and A. Avila-de Melo. 2017. Dynamic of the Flemish Cap commercial stocks: use of a Gadget multispecies model to determine the relevance and synergies among predation, recruitment, and fishing. Canadian Journal of Fisheries and Aquatic Sciences 74:582-597.

Rindorf, A., Cardinale, M., Shephard, S., De Oliveira, J. A., Hjorleifsson, E., Kempf, A., ... & Simmonds, J. (2017). Fishing for MSY: using "pretty good yield" ranges without impairing recruitment. ICES Journal of Marine Science, 74(2), 525-534.

SEDAR. 2020a. SEDAR 69 – Atlantic Menhaden Benchmark Stock Assessment Report. SEDAR, North Charleston SC. 691 pp. available online at: http://sedarweb.org/sedar-69

SEDAR. 2020b. SEDAR 69 – Atlantic Menhaden Ecological Reference Points Stock Assess-ment Report. SEDAR, North Charleston SC. 560 pp. available online at: http://sedarweb.org/sedar-69

Sparholt, H., Bogstad, B., Christensen, V., Collie, J., vanGemert, R., Hilborn, R., Horbowy, J., Howell, D., Melnychuk, M.C., Pedersen,S.A., Sparrevohn, C.R., Stefansson, G., and Steingrund, P. 2020. Estimating Fmsy from an ensemble of data sources to account for density dependence in Northeast Atlantic fish stocks. ICES Journal of Marine Science, doi:10.1093/icesjms/fsaa175. Published online 15 October 2020.

Tableau, A., Collie, J.S., Bell, R.B. and Minto, C. 2019. Decadal changes in the productivity of New England fish populations. Canadian Journal of Fisheries and Aquatic Sciences 76(9): 1528-1540. https://doi.org/10.1139/cjfas-2018-0255

Thorpe, R. B. & De Oliveira, J.A.A. 2019. Comparing conceptual frameworks for a fish community MSY (FCMSY) using management strategy evaluation: an example from the North Sea. ICES Journal of Marine Science 76(4): 813-823.

Thorson, J. T., Cope, J. M., Branch, T. A., & Jensen, O. P. (2012). Spawning biomass reference points for exploited marine fishes, incorporating taxonomic and body size information. Canadian Journal of Fisheries and Aquatic Sciences, 69(9), 1556-1568.

Uphoff Jr, J. H. and Sharov, A. (2018). Striped Bass and Atlantic Menhaden Predator–Prey Dynamics: Model Choice Makes the Difference. Mar. Coast. Fish. 10(4), 370-385.

Van Deurs, M., Brooks, M.E., Lindegren, M., Henriksen, O. and Rindorf, A. (2020) Biomass limit reference points are sensitive to estimation method, time-series length and stock development. Fish and Fisheries, DOI: 10.1111/faf.12503

Walters, C.J. 1987. Nonstationarity of production relationships in exploited populations. Can. J. Fish. Aquat. Sci. 44(Supplement 2), 156-165.

Zhang, F., Regular, P., Wheeland, L., Rideout, R., Morgan, M.J. (2020). Accounting for non-stationary stock-recruitment relationships in the development of MSY-based reference points. ICES Journal of Marine Science. doi: 10.1093/icesjms/fsaa176

Zhang, F., Reid, K. B., & Nudds, T. D. (2018). Ecosystem change and decadal variation in stock-recruitment relationships of Lake Erie yellow perch (*Perca flavescens*). ICES Journal of Marine Science, 75(2), 531–540. doi: 10.1093/icesjms/fsx188

Annex 1: List of participants

Name	Institute	Country	Email
Eider Andonegi	AZTI-Tecnalia	Spain	eandonegi@azti.es
Per Arneberg	Institute of Marine Research	Norway	per.arneberg@hi.no
Valerio Bartolino	Department of Aquatic Resources-SLU Aqua - Swedish University of Agricultural Sciences	Sweden	valerio.bartolino@slu.se
Aaron M. Berger	NOAA Fisheries - National Oceanic and Atmospheric Administration	United States	aaron.berger@noaa.gov
Bjarte Bogstad	Institute of Marine Research	Norway	bjarte.bogstad@hi.no
Lisa Borges	Fishfix, Lisbon	Portugal	info@fishfix.eu
Noel Cadigan	Fisheries and Marine Institute - Memorial University of Newfoundland	Canada	noel.cadigan@mi.mun.ca
Massimiliano Cardinale	Department of Aquatic Resources-SLU Aqua - Swedish University of Agricultural Sciences	Sweden	massimiliano.cardinale@slu.se
Jeremy Collie	Graduate School of Oceanography - University of Rhode Island	United States	jcollie@uri.edu
Mikael van Deurs	DTU Aqua, National Institute of Aquatic Resources	Denmark	mvd@aqua.dtu.dk
Sarah Gaichas	Northeast Fisheries Science Center - NOAA Fisheries	United States	sarah.gaichas@noaa.gov
Jenni Grossmann	ClientEarth	United Kingdom	jgrossmann@clientearth.org
Jan Horbowy	National Marine Fisheries Research Institute	Poland	horbowy@mir.gdynia.pl
Daniel Howell	Institute of Marine Research	Norway	daniel.howell@hi.no
Melissa A. Karp	Office of Science and Technology - NOAA Fisheries	United States	melissa.karp@noaa.gov
Mathieu Lundy	Agri-food and Biosciences Institute	United Kingdom	mathieu.lundy@afbini.gov.uk
Wendy E. Morrison	NOAA Fisheries - National Oceanic and Atmospheric Administration	United States	wendy.morrison@noaa.gov
Martin Pastoors	Pelagic Freezer-Trawler Association	Netherlands	mpastoors@pelagicfish.eu
Alfonso Pérez Rodriguez	Institute of Marine Research	Norway	alfonso.perez.rodriguez@hi.no
Jay Peterson	Office of Science and Technology - NOAA Fisheries	United States	jay.peterson@noaa.gov
David Reid	Marine Institute	Ireland	david.reid@marine.ie
Anna Rindorf	DTU Aqua, National Institute of Aquatic Resources	Denmark	ar@aqua.dtu.dk
Amy Schueller	NOAA Fisheries - National Oceanic and Atmospheric Administration	United States	amy.schueller@noaa.gov
Henrik Sparholt	University of Copenhagen	Denmark	henrik.sparholt@gmail.com
Vanessa Trijoulet	DTU Aqua, National Institute of Aquatic Resources	Denmark	vtri@aqua.dtu.dk
John Woodlock	Irish Seal Sanctuary	Ireland	jwoodlock@eircom.net
Fan Zhang	Fisheries and Marine Institute - Memorial University of Newfoundland	Canada	fan.zhang@mi.mun.ca