ICES AD HOC REPORT 2018

ICES ADVISORY COMMITTEE

ICES CM 2018/ACOM:60

REF. ACOM

Report on IUCN assessments and fisheries management approaches

2018

Authored by Sarah Millar and Mark Dickey-Collas

International Council for the Exploration of the Sea
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

Recommended format for purposes of citation:
Millar, S., and Dickey-Collas, M. 2018. Report on IUCN assessments and fisheries management approaches. ICES CM 2018/ACOM:60. 109 pp. https://doi.org/10.17895/ices.pub.4676

For permission to reproduce material from this publication, please apply to the material in this report may be reused using the recommended citation. ICES may only grant usage rights of information, data, images, graphs, etc. of which it has ownership. For other third-party material cited in this report, you must contact the original copyright holder for permission. For citation of datasets or use of data to be included in other databases, please refer to the latest ICES data policy on ICES website. All extracts must be acknowledged. For other reproduction requests please contact the General Secretary.

The document is a report of an Expert Group under the auspices of the International Council for the Exploration of the Sea and does not necessarily represent the views of the Council.

© 2018 International Council for the Exploration of the Sea

Contents

Executive Summary

ICES received a request from the EU to compare the IUCN methods for assessing conservation status of species with the assessment methods used by fisheries management. ICES was also asked to explore the pertinence of IUCN methods to fisheries management challenges. This report was prepared by the ICES secretariat with contributions from two studies and consultation with IUCN researchers, ACOM and FAO.

The report provides the background to the issue, it reviews previous published studies, it notes the similarities between the two approaches and describes the previous ICES advice (from 2009 and 2013) on IUCN/fisheries issues. Both the IUCN Red List and fisheries management approaches develop and use an evidence base to inform society about the conservation status of mature organisms. IUCN focuses on all species and only conservation status, while fisheries management works on the units of marine organisms that are exploited and introduces the additional concept of sustainable exploitation. Both require transparency in decision making, a clear framework for the process of assessment, the use of best available science, robust scrutiny and review, regular updates as situations change or new knowledge is developed. Both are globally accepted methods for their respective objectives.

Against the background of differing rationale for their methods, the report suggests that the main differences between the two approaches are:

- the attitude to the evidence base (linked to different interpretations of the application of the precautionary principle and application of methods),
- the use of reference points for assessing the state of the population contrasting to the use of recent trends in population abundance (reflecting difference in objectives)
- the acceptance of projections to assess the risk of proposed management actions (the applicability of fisheries management strategy evaluation (MSE) as a tool to assess future options and test extinction risk, i.e. criterion E).

There are also minor differences such as stocks being assessed rather than species, the frequency of data collection and the impact of fishing on generation time.

A exploration of the IUCN and CFP classification of EU stocks and found the majority of stocks were similarly assigned (72%) when comparing biomass estimates, but the similarity dropped to 55% when comparing estimated of fishing mortality to conservation status. There was no bias by IUCN methods towards exaggerating the threat status. The number of misses (also referred to as false positives, that is, when stocks are assessed as being below biomass reference points but are not given a threatened categorisation by the IUCN method) was 23% for biomass and 40% for fishing mortality suggesting that the IUCN approach does not provide the full evidence to deliver CFP objectives.

A comparison between IUCN categories of conservation status and fisheries reference points using management strategy evaluation on four simulated stocks with varying life history characteristics and data quality, highlighted some differences. Stocks at low stable abundance will not alert the IUCN categories but will require action, to achieve the CFP objective of maintaining populations of harvested species above levels which can produce the maximum sustainable yield. The analysis showed that a harvest control rule is successful at maintaining stocks above critical biomass limits, even in data-poor cases, and extinction risk is <1%.

The report concludes that IUCN listing and fisheries management approaches have different but complementary objectives and these objectives result in differing interpretation of the precautionary principle, and differing attitudes towards data and the knowledge base. The use of IUCN criterion A to assess marine fish focuses on the direction and magnitude of the recent trend in biomass. This operates well to alert management authorities to recent change. It could be considered the equivalent of a warning-bell in the context of data deficient, commercially exploited fish and is likely a crucial alarm for non-commercially exploited species.

The IUCN approach cannot advise on maximising yield for fisheries and assessing risk to fish stocks in relation to biomass reference points. It cannot detect that a stock is over-fished if the biomass is very low and stable. It could be argued that when a fish stock is exploited, the IUCN approach can underestimate the risk to a stock being outside safe biological limits.

IUCN has the objective to list all species. For commercially exploited fish, this can be easily achieved by coupling an IUCN assessment to current stock assessments. This would supplement the knowledge base to assessing the CFP objectives of long term sustainability and minimising negative impacts of fishing activities. This would aid Europe in assessing progress towards the Aichi targets and UN SDG 14.

The approach used in Norway of combining fisheries and IUCN approaches builds a robust warning system for marine fish. If used alone, IUCN approaches would not be reliable as an evidence base to deliver the objectives of the CFP, they would need to be supplemented with approaches that consider the status on the biomass relative to reference points and some consideration of the exploitation/bycatch rate.

1 Introduction

1.1 The request

In November 2017, ICES received the following request from the European Commission (DGMARE):

Analysis of the IUCN process for the assessment of the conservation status of marine species in comparison to the process used by fisheries management bodies.

The key objectives of the work were:

- Perform a critical analysis of the IUCN assessment process for marine species in comparison the scientific process of relevant fisheries organisations (in particular ICES, scientific committees of RFMOs); including how data/information from different sources used by IUCN are weighted to derive a final perception of decline
- Provide an analysis of the criteria used by IUCN and their suitability for marine fish species
- Provide a comparison between IUCN categories of conservation status and fisheries reference points for defining sustainability at both the species and stock level
- Identify pros and cons of the two processes
- Compare recent IUCN listings for key EU stocks with corresponding information from stock assessments.
- Case studies species should cover a range of taxonomic levels and also include a range of data rich and data limited stocks. Case study species could usefully be agreed between ICES and MARE prior to any workshop.
- Assess suitability and reliability, and advantages of disadvantages of incorporating IUCN assessments for fisheries management and other relevant decision making processes.

Supporting information: International Union for Conservation of Nature IUCN assessments are increasingly used as source of "scientific" information for the perception of the stock status of marine fish species, including those that are falling under the purview of fisheries management bodies. These assessments are becoming increasingly influential in shaping the public opinion on conservation issues, as well as, in supporting decisions in national and international organisations, in particular multilateral environmental agreements. However, there are also increasingly examples of inconsistencies between the outcome of the IUCN assessments and those of scientific bodies supporting fisheries management organisations, which indicate that there is a need to better understand the pros and cons of the two approaches and most importantly their reliability in the context of decision making processes for the conservation and management of marine fish species. In addition, it is not clear how this process works in terms of methodology, reproducibility, accuracy, coherence and comparability of outputs with fisheries assessments etc.

As instructed, the request was further discussed via telephone and email with members of the European Commission and the proposed approach was circulated across DGMARE for feedback in the spring of 2018.

1.2 Fisheries management objectives

A context is required when comparing IUCN methods for assessing the conservation status of a species with fisheries management approaches for stocks. For the purposes of this report, the lead objectives of the reformed common fisheries policy (CFP, EU 2013, article 2) will be used, and the policy lists the first three EU fisheries management objectives as:

1. *The CFP shall ensure that fishing and aquaculture activities are environmentally sustainable in the long-term and are managed in a way that is consistent with the objectives of achieving economic, social and employment benefits, and of contributing to the availability of food supplies.*
2. *The CFP shall apply the precautionary approach to fisheries management, and shall aim to ensure that exploitation of living marine biological resources restores and maintains populations of harvested species above levels which can produce the maximum sustainable yield.*
3. *The CFP shall implement the ecosystem-based approach to fisheries management so as to ensure that negative impacts of fishing activities on the marine ecosystem are minimised, and shall endeavour to ensure that aquaculture and fisheries activities avoid the degradation of the marine environment.*

This report will assume that the rationale for the comparison of methods will be the following key phrases:

- fishing … activities are environmentally sustainable in the long-term
- apply the precautionary approach
- maintains populations of harvested species above levels which can produce the maximum sustainable yield
- negative impacts of fishing activities on the marine ecosystem are minimised
- fisheries activities avoid the degradation of the marine environment

2 Methods

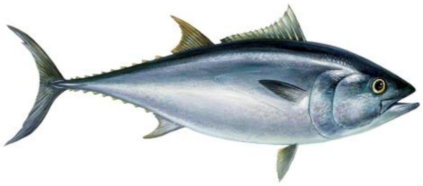
ACOM agreed to answer the request in the following manner:

- A literature study and broader discussions with IUCN, FAO and CITES were to be carried out by the ICES secretariat.
- A comparison of the classification of EU stocks (ICES and ICCAT) through the IUCN process and fisheries management (CFP) approaches was contracted out to the University of Aberdeen. The choice of stocks for comparison was agreed with DGMARE and it was agreed not to cover non-ICCAT stocks in the Mediterranean.
- A comparison between IUCN categories of conservation status and fisheries reference points for defining sustainability at both the species and stock level was contracted out to the University of Massachusetts. This was to be done through simulation studies of 4 case study species representing a range of population characteristics and veracities of data provision. The stocks for the case studies were agreed with DGMARE as Blue fin tuna (*Thunnus thynnus*), Blue shark (*Prionace glauca*), Atlantic cod (*Gadus morhua*), Red (blackspot) seabream (*Pagellus bogaraveo*) (see Box 1).
- Independent external review of the report and studies.
- Based on these activities, a draft advice was to be drawn up by the ICES secretariat which would form the underlying basis for an advice drafting group which is due to meet toward the end of 2018.
- Formal ACOM review and release of the advice, November 2018.

Early in the process, ACOM was also asked to advise of the appropriate biomass reference point for the comparison of classification of EU stocks by IUCN and fisheries management status. ACOM agreed to use B_{pa}, the precautionary biomass reference point for comparison with the range of categories used by IUCN.

Box 1. Species used as case studies comparison between IUCN categories of conservation status and fisheries reference points for defining sustainability.

1. **Blue fin tuna** (*Thunnus thynnus*) - typical emblematic large pelagic, assessed by ICCAT. Base the simulations on population similar to western Atlantic stock. Has been subject of CITES and FAO considerations http://www.fao.org/docrep/013/i1899e/i1899e00.htm. Listed as endangered by IUCN http://www.iucnredlist.org/details/21860/0



2. **Blue shark** (*Prionace glauca*) – relatively data poor pelagic elasmobranch, assessed by ICCAT and other stocks around the world. Stock assessments mostly based on CPUE and catches using SS3. Listed as near threatened by IUCN and not on the CITES list. http://www.iucnredlist.org/details/39381/0

3. **Atlantic cod** (*Gadus morhua*), typical emblematic gadoid, data rich. Stock assessments using survey and age data with many different models. Useful as has been a target for IUCN concerns (http://www.ices.dk/sites/pub/Publication%20Reports/Advice/2013/Special%20requests/Helcom_Baltic_Cod_Red_List.pdf). Listed as vulnerable by IUCN http://www.iucnredlist.org/details/8784/0

4. **Red (blackspot) seabream** (*Pagellus bogaraveo*). Data poor, deep sea species of commercial interest. A useful example of a species that is representative of data poor deep sea issues. In some areas no evidence of recovery after fishing. ICES assesses three stocks- Azores, Iberian Atlantic and Bay of Biscay Celtic seas.

   - http://ices.dk/sites/pub/Publication%20Reports/Advice/2017/Special_requests/EU_sbr-x_review.pdf
   - http://ices.dk/sites/pub/Publication%20Reports/Advice/2016/2016/sbr-678.pdf
   - http://ices.dk/sites/pub/Publication%20Reports/Advice/2016/2016/sbr-ix.pdf

   Listed as near threatened by IUCN http://www.iucnredlist.org/details/170244/0

3 Published discussion about comparison of IUCN and fisheries management approaches.

3.1 Introduction

The IUCN red list is a global compendium of the conservation status of species. It can be seen as an indicator that informs and catalyses action, tracks the progress of conservation efforts (including towards the Aichi biodiversity targets), and communicates urgency for conservation issues. The process and methodology to assess species is well described and is based on empirical, science derived evidence. The IUCN Red List works under the precautionary principle.

The methods used under the fisheries management umbrella are tools which directly explore and influence management decisions. They are also science based, and include surveys, stock assessments and management strategy evaluations. Management Strategy Evaluation (MSE) involves a simulated fish stock, the dynamics of which are known, on which potential harvest control rules can then be tested to see how stocks respond to management. The precautionary approach and the incorporation of ecosystem considerations are at the forefront of fisheries management.

There is a perception that the two systems can still seem at odds to each other in their assessment of the vulnerability of species to extinction. Part of this disconnect may come from the different objectives that are treated in a similar manner: the IUCN Red List assesses the vulnerability of a species to extinction risk, whilst fisheries management and stock assessment typically work to manage stocks under Maximum Sustainable Yield (MSY), which usually implies a population size below that of an unharvested population and focuses on sustainable exploitation with biomass targets (or limit reference points). There is also a suggestion that in data deficient cases, IUCN approaches could be used as a tool for fisheries management (Dulvy *et al.*, 2005).

Previous published work to evaluate the compatibility of the two systems has either tested the reliability of IUCN criterion A (population decline) using a population dynamic simulation model to estimate the probability of extinction and compared this to the IUCN Red List, or compared current stock assessments with the IUCN Red List categorisation (e.g. Powles *et al.*, 2000; Punt, 2000; Dulvy *et al.*, 2005; Fernandes *et al.* 2017).

Marine species now represent 15% of all species assessed with the IUCN Red List methodology (IUCN, 2017). The number of extinctions in the marine environment has been remarkably low, as reported by Poles *et al.* (2000) and by McCauley *et al.* (2015) where they examined marine extinctions over the past 514 years. A small fraction were found to have a threat categorisation, although many species assessed were found to be data deficient. The small number of extinctions (15 in the past 514 years and none in the last 50) was attributed to the wider distributional range of marine species compared to their terrestrial counterparts, their lower endemism and higher dispersal. However, size spectra of many communities had been found to have changed over the time period, and 90% of large pelagic fish show a range contraction.

Additionally, Fernandes *et al.* (2017) found that in their analysis the commercial exploitation of a marine fish species was *not* one of the most significant factors in determining extinction risk. The authors were clear, that their results did not suggest fishing is unimportant as a risk to extinction, but that life history traits such as body size and time to reach maturity had the greatest impact on how vulnerable a species would be to extinction.

3.2 Comparability

Since 1994, IUCN has made efforts to work transparently to determining risks of extinction applicable to all species, using five quantitative criteria. The 1996 IUCN Red list of Animals included over 100 marine fishes (Hudson and Mace, 1996). Papers questioning the suitability of the IUCN criteria for determining threat of extinction, were prevalent at the end of the 1990s. Musick (1998), Hudson (1999) and Powles *et al*. (2000) expressed concerns that the IUCN listings did not reflect the true extinction risk, especially if the decline was a result of management action focusing on MSY. By using criterion A (population decline) almost exclusively to assess extinction-risk, the reduction in stock biomass/abundance evident in most commercial fisheries at the commencement of fishing would qualify the species for a threated status. In addition, the Powles *et al*. (2000) paper featured an incorrect and regularly-referenced figure that erroneously implied a fish stock/species would only get a threatened listing by IUCN when the stock had been completely collapsed by management, which is not the case.

Punt (2000) used a stochastic population dynamic model to examine the magnitude of fishing mortality (F) at which a population is driven to extinction (F_{crash}). Results showed that there was a substantial probability of incorrectly identifying species being harvested at F_{MSY} as threatened during the initial fishing down phase, and of not identifying species that would be at risk of extinction if F was not reduced.

Since publication of the papers by Musick, Punt, and Powles *et al*, criteria A has been updated, and better encompasses the dynamics of commercially exploited fish populations, meaning that much of the simulation work from these papers, and their conclusions, are no longer valid. Criteria A1 is used for cases where a population decline occurred in the past but the decline has ceased, is clearly understood and is reversible (IUCN, 2012). In 2001, IUCN updated their criterion A to the following:

A. Reduction in population size based on any of the following:

1) An observed, estimated, inferred or suspected population size reduction of ≥ 90% over the last 10 years or three generations, whichever is the longer, where the causes of the reduction are clearly reversible AND understood AND ceased, based on (and specifying) any of the following:
   - a) direct observation
   - b) an index of abundance appropriate to the taxon
   - c) a decline in area of occupancy, extent of occurrence and/or quality of habitat
   - d) actual or potential levels of exploitation
   - e) the effects of introduced taxa, hybridization, pathogens, pollutants, competitors or parasites.

2) An observed, estimated, inferred or suspected population size reduction of ≥ 80% over the last 10 years or three generations, whichever is the longer, where the reduction or its causes may not have ceased OR may not be understood OR may not be reversible, based on (and specifying) any of (a) to (e) under A1.

3) A population size reduction of ≥ 80%, projected or suspected to be met within the next 10 years or three generations, whichever is the longer (up to a maximum of 100 years), based on (and specifying) any of (b) to (e) under A1.

4) An observed, estimated, inferred, projected or suspected population size reduction of ≥ 80% over any 10 year or three generation period, whichever is longer (up to a maximum of 100 years in the future), where the time period must include both the past and the future, and where the reduction or its causes may not have ceased OR may not be understood OR may not be reversible, based on (and specifying) any of (a) to (e) under A1.

Criteria A has been almost universally used for marine fishes due to challenges caused by estimating the distribution area of these species which would allow of criteria B. Criteria A has been assessed using simulations that project forward up to 500 years using the IUCN definitions of extinction risk (between 50–90% chance of going extinction in the next 10 years or three generations, whichever is shorter reference?).

In 2005, Dulvy *et al.* authored one of the first and most influential papers that compared the new IUCN criteria A1 (see above) to stock assessment classifications, as well as the IUCN criteria E (quantitative analysis of extinction risk), and found that the two were largely compatible. Dulvy *et al*. (2005) compared quantitative stock assessments to IUCN Red Listing. They described their results using the terms "hits" (Red List agrees with the stock assessment), "misses" (Red list does not list a species as threatened but the stock assessment shows stocks is not exploited unsustainably) and "false alarms" (Red list assesses as a species as threatened but the stock assessment shows stock is being exploited sustainably). In this analysis 0 stocks were "false alarms", 48% of stocsk were"misses" when comparing criterion A and biomass reference points, and 25% "misses" when comparing criterion E with biomass reference points. The new decline rate criteria provided risk categorisation consistent with population viability when applied to exploited marine stocks. They also noted IUCN E criteria has been applied to fewer stocks that the criteria A (that can also be used for data deficient stocks).

Rice and Legace (2007) undertook work similar to that of Dulvy *et al*., but increased the number of species from 76 to 89, and used 35 years of population data. The authors found that assessing extinction risk using the IUCN criteria B (geographic range decline) and criteria C (number of mature individuals) were in harmony with limit reference points used by fisheries management agencies: fishing mortality would have been reduced to zero by managers well before there was a chance of extinction. However, an important finding of this study showed that there was a large potential for conflict between fisheries and risk-of-extinction approaches when considering the extent of population declines, that is, criterion A, which remains the criterion by which most fish species are categorised.

Davies and Baum (2012) compared IUCN and fisheries metrics and concluded that biological reference points for commercially exploited species and IUCN analysis were compatible and well aligned, despite the "mathematical disconnect". They concluded that the Red List was not biased towards exaggerating the threat status; that only the question of management responses for populations of mutual concern were left unresolved. It is also of interest that of the species considered, whereas 13.5% of Red Listed marine fishes were classed as threatened, 40% of populations with a stock assessment were classed as being below biomass limit reference points, indicating the biomass limit reference points tend to be more conservative than the IUCN criteria for threatened.

More recently Fernandes *et al*. (2017) compared the IUCN Red List of European fish with 114 stock assessments (of 31) species and found that no species classified by IUCN as threatened were considered to be unsustainably harvested by fisheries management agencies. The paper concluded that the stock assessments carried out by ICES and the IUCN Red Listing analysis should be seen as "complementary classification schemes".

Both Davies *et al.* (2012) and Fernandes *et al.* (2017) found that certain life history characteristics contribute to which species are at the greatest risks of extinction. IUCN listings reflect this in the productivity of the species (where productivity is defined as the maximum per capita growth of a population).

There is an assumption in all these methods of comparison that the stock assessment represents the true state of the system; in reality, even if both IUCN listing and stock assessment class a species as threatened/not threatened, they may both be incorrect. Simulation work, such as that undertaken by Fay (Annex 2, this report) manages to circumvent this problem as the true starting population is known.

3.3 Timing

Once a species has been assessed, IUCN aim to update the Red List assessment every 10 years. As Rondinini *et al.* (2013) point out, however, whilst the IUCN plan to increase the number of species covered by the Red List, there is currently no operational plan to manage these updates, and as of 2013, 17% of assessments were out of date (Rondinini *et al*., 2013). IUCN use relevant data from multiple sources including publications, grey literature, and expert knowledge. Since 2000, the assessments have been conducted via physical workshops with participating experts. This scientific rigour has an associated cost that has an impact on the frequency of IUCN updates.

Stock assessments are conducted by organisations such as ICES and ICAAT on a more frequent basis. Typically this is between every one to five years (see latest ICES published advice).

Investment in both stock assessments and IUCN assessment is large. Norway is an example of a country investing large amounts time and resources to ensure its stocks do not receive a threatened listing. The Norwegian focus on fisheries management in recent years has been on rebuilding the commercially important stocks through the Ecosystem Approach to Fisheries Management (EAFM). The official Norwegian Red List for Species has become an important tool, or yardstick for management of economically important species. The aim is to minimise the risk of future listing of a species, and management measures are tailored for species already listed in order for that species to be removed from the list. The "Stock Table" and "Fisheries table" that inform the Red List are updated each spring, although at present only species with an annual catch of more than 100 tonnes, and species on the official Red List, are subject to evaluation. The state of all species including the data-poor ones are assessed according to the IUCN criteria every 5 years (Gullestad *et al*., 2017).

3.4 Summary

The most recent publications comparing the IUCN Red List with stock assessments undertaken by fisheries management agencies show the two as complementary products that offer differing insights and levels of information, and should not necessarily be taken as directly comparable but could be complementary.

+

4 Similarities between IUCN and fisheries management approaches

The IUCN Red List of Threatened Species is widely recognized as the most comprehensive, objective global approach for evaluating the conservation status of plant and animal species (Rondinini *et al*., 2014). The IUCN Red List has grown in size and complexity and now plays an increasingly prominent role in guiding conservation activities of governments, NGOs and scientific institutions (Rice and Legacè, 2007; Gullestad *et al*., 2017). The introduction in 1994 of a scientifically rigorous approach to determine risks of extinction that is applicable to all species, is now considered a world standard. The IUCN Red List is underpinned by information management tools (the Species Information Service) which facilitates the collection, management and processing of species data from workshop to publication on The IUCN Red List (http://www.iucnredlist.org/about/overview).

The goal of The IUCN Red List is to provide information and analyses on the status, trends and threats to species in order to inform and catalyse action for biodiversity conservation.

The IUCN Red List aims to:

- Establish a baseline from which to monitor the change in status of species;
- Provide a global context for the establishment of conservation priorities at the local level;
- Monitor, on a continuing basis, the status of a representative selection of species (as biodiversity indicators) that cover all the major ecosystems of the world.

Fishery management systems in many regions rely on regularly scheduled stock assessments to determine stock status and provide management advice for achieving fishery and conservation objectives. Stock assessments typically involve the synthesis of information on life history, fishery monitoring, and resource surveys for estimating stock size and harvest rate relative to sustainable reference points. Stock assessments also frequently involves forecasting the response of the resource to alternative management scenarios (Hilborn and Walters, 1992; Quinn and Deriso, 1999). Stock assessments are usually carried out by applying mathematical models that are fitted to available information to provide simplified representations of population and fishery dynamics (Cadrin and Dickey-Collas, 2015). Over the last century, stock assessment methods have progressed from descriptive models, often assuming equilibrium, to elaborate statistical models with many estimated parameters and formal approaches to evaluating uncertainty (Cadrin and Dickey-Collas, 2015). The combination of stock assessments and management strategy evaluations are used to explore the consequences of action in relation to fisheries management objectives (see chapter 1 for the lead objectives of the EU CFP).

Both the IUCN Red List and fisheries management approaches develop and use an evidence base that informs society about the conservation status of mature organisms. The IUCN focuses on all species and only conservation status, while fisheries management works on the units of marine organisms that are exploited and introduces the additional concept of sustainable exploitation.

Both require transparency in decision making, a clear framework for the process of assessment, the use of best available science, robust scrutiny and review, regular updates as situations change or new knowledge is developed. Both are globally accepted methods for their respective objectives.

5 Previous ICES advice related to IUCN listing

5.1 Request from Norway 2009

In 2009, ACOM was asked by Norway to provide advice on the IUCN listing of marine fish species (ICES, 2009a). The advice was based on an exploration of Population Viability Analysis (PVA) methods to assess extinction risk for commercial exploited fish stocks (WKPOOR1; ICES, 2009b) and the application of those methods to cod and redfish stocks (WKPOOR2; ICES, 2009c). ICES stated:

"There are three general methods for evaluating extinction risk: (1) screening methods, such as the IUCN redlisting criteria; (2) simple population viability analysis based on time trends; and (3) age structured population viability analysis. The rate of false positives (prediction of extinction which does not occur) and false negatives (the occurrence of unpredicted extinction) is likely to be the highest for screening methods, lower for simple population viability analysis based on time trends, and lowest for age structure population viability analysis. None of the methods are considered reliable for accurately estimating the probability of extinction, but they may be useful to evaluate the relative probability of extinction between species or between management options."

Later in the advice ICES stated:

"Screening methods may be useful to prompt a more comprehensive analysis, but should not be used as the basis for a listing decision when more detailed data are available, as is typically the case for exploited marine species. Screening methods also only provide an evaluation of stock status at a point in time. They do not include a projection into the future which is more useful for estimating the probability of extinction. As well, criteria based on the rate or magnitude of population decline may overlook the fact that even well managed exploited fish populations can experience large declines. Furthermore, in some cases even a small additional decline may induce a population to pass a tipping point and lead to an increased chance of extinction.

Population Viability Analysis (PVA) is a method that projects a population forward in time using uncertainty to make statements about the probability of population abundance falling below some predetermined level in a given number of years. PVA is useful to indicate the relative risk of extinction (e.g., between stocks) rather than to estimate the absolute probability. The PVA is a forecast of what would be likely to happen to a stock if current conditions remain unchanged throughout the projection period. This assumption of stationarity implies that the conditions that generated the observed values will continue into the future.

Another approach is the Age Structured Population Viability Analysis. In the standard application of this approach the simple PVA is augmented to account for life stage/age structure allowing density dependence and other forms of non-stationarity to occur in the projections. This approach allows comparison of the relative probability of extinction for alternative management options.

Standard fishery models can also be used to examine the risk of extinction. Stock and recruitment estimates can be compared to the replacement line under the current mortality rate. When total mortality is too high, the replacement line will be to the left of recruitment values associated with low stock size, causing the stock to decline. If depensation is present in the stock-recruitment relationship (or if the stock-recruitment relationship changes over time causing a smaller slope at the origin), too high a mortality rate will cause the stock to eventually go extinct. There is no time period involved in this approach, but continued recruitment below the replacement line [at low stock size] *implies a high probability of extinction."*

5.2 Request from HELCOM 2013

In February 2013, ICES published advice on a review of The Baltic Marine Environment Protection Commission (HELCOM) draft Red List assessment of cod (*Gadus morhua*, ICES 2013). HELCOM is an intergovernmental organisation governing the Convention on the Protection of the Marine Environment of the Baltic Sea Area. While this request was specific to the HELCOM IUCN assessment of Baltic cod, some issues can be viewed as relevant to the current request from DGMARE.

The advice answered the following questions:

Has the HELCOM Red List assessment been carried out appropriately following the criteria of IUCN?

> *ICES advises that Criterion E should be used to assess reduction in population size, rather than screening assessments using Criterion A for the three cod stocks in the Baltic. In the current HELCOM assessment, ICES advises that Subcriterion A1 should have been used rather than Subcriterion A2 except for the Kattegat stock, and that habitat loss should have been assessed using Criterion B rather than Criterion A, because spawning-stock biomass (SSB) trends were available.*

> *ICES advises that due to the separate past (and future) trajectories of the three stocks, it is not appropriate to assign one IUCN category collectively for all cod in the Baltic.*

Has the assessment utilized correctly all appropriate data on the development of the cod stock(s) and its habitats?

> *ICES advises that further data should have been used in the assessment. With regards to habitats, the best information was generally used, but interpreted in an inconsistent manner, and should have been assessed using Criterion B.*

Has the generation time of cod been estimated properly?

> *ICES advises that the calculation of generation time of cod was consistent with IUCN guidelines, but some of the parameters used were inappropriately specified.*

Does any significant immigration exists between the Baltic Sea stock(s) and the North Sea population?

> *There is insufficient information to advise on the degree of mixing between the cod stocks in the Baltic and in the North Sea. From an assessment and management perspective, separation is assumed between the stocks.*

The advice also stated "*ICES views a stock assessment with a projection as an appropriate analysis of the likely extinction risk of a commercially exploited marine organism due to reduction in population size. This is a more effective tool than the IUCN red listing Criterion A*."

6 Differences in approaches

This section explores the distinctions between the approaches used for assessment by IUCN and fisheries management.

6.1 Evidentiary attitude and the precautionary approach

There is a difference between the approach to the evidence base and use of methods between IUCN and fisheries management. IUCN request that as many criteria as possible be used to assess the status of a taxa. The criteria that produces the highest threat category should be the method used for the assessment. This is approach has been adopted in support of the IUCN understanding of the precautionary principle.

The IUCN guidelines state that assessors should resist an evidentiary attitude and adopt a precautionary but realistic attitude to uncertainty when applying the criteria (IUCN 2017, section 3.2.3). "Evidentiary attitude" is described as using a threatened category only when strong evidence exists. This is further described as better to use "plausible lower bounds rather than best estimates" (IUCN 2017, section 3.2.4).

This does not however judge which criteria is most suitable considering the evidence base, as done in fisheries management. The criteria with the worst assessment might also be the criteria with the least reliable evidence base (see ICES 2009, "*The rate of false positives (prediction of extinction which does not occur) and false negatives (the occurrence of unpredicted extinction) is likely to be the highest for screening methods*"). This also does not encourage improvement of the evidence base for assessment. The use of criterion E might lead to the most robust assessment of extinction threat, but the extra research may not be warranted if a higher, but less certain, extinction risk category is highlighted using for example criteria A or B.

There are other consequences of this precautionary versus evidentiary attitude such as providing higher discrimination ability among taxa and providing more meaningful conservation information using the evidentiary attitude (Romeiras *et al*., 2016).

6.2 Stock verses species

The difference between stock and species is regularly cited as a big difference between the two approaches (Musick, 1999; Davies and Baum, 2012; Fernandes *et al*., 2017). Stocks are a spatially disaggregated component of the species distribution and recognition of the stock composition and distribution of a species is an important aspect of fisheries management. IUCN allow for regional assessments, and the delineation of fish stocks has evolved from an original pragmatic interpretation of exploitation and biological information (Reiss *et al*., 2009; IUCN, 2017b) and current IUCN practice is that the assessed unit must be a biological unit not "defined by political or national boundaries" (IUCN, 2017, section 2.1.1). This approach holds true for fish stocks too. There is extensive guidance from IUCN on how to merge signals from differing populations of the same species (that can be used for fish stocks in an equivalent way). It should be noted that before carrying out stock based assessments under IUCN,, the assessor must conduct a full species level assessment (IUCN, 2017, section 2.1.1).

6.3 Biomass reference points, depletion and trends

The different operational purposes of IUCN and fisheries management mean that issues such as the level of a population and trends in the assessments are treated differently: "*It must be remembered … that the IUCN Red List Criteria are designed to identify taxa* *that exhibit symptoms of endangerment, and not simply depletion or conservation priority*" (IUCN 2017, section 5.5). In practice this means that a fish that is at a low stable level and could be below biological safe reference points (not only overfished but below biomass limit reference points, thus in danger of limiting recruitment) in a fisheries management context could be classes as least concern under IUCN. The EU has the following objective for the common fisheries policy; "…*maintains populations of harvested species above levels which can produce the maximum sustainable yield*" (EU 2013, Article 2). This thus imbeds biomass reference points and assessments of state as a requirement of fisheries management. So, while acknowledging and accounting for trends, fisheries management suggests action based on levels in relation to reference points. This means that contradictory assessments can appear to be made between IUCN and fisheries approaches.

IUCN acknowledge that the nature of fishing will reduce the biomass of a fish stock and expects that with time the level will stabilise at a new level. Any initial exploitation of a stock should trigger IUCN criterion A (IUCN 2017, section 5.6.1). IUCN emphasise that once well managed, there should be no decline rate (a historical reduction followed by long-term stabilization, IUCN, 2017, section 5.5).

In the context of ICES advice rules, IUCN is less precautionary in that it assumes that fishing at F/F_{MSY} <1 does not threaten the biomass. The ICES advice rule acknowledges that for some stocks of fish, the biomass can fall below B_{lim} even when fishing at or below F_{MSY}. The ICES advice rule requires that F_{MSY} be reduced if there is a >5% chance of the biomass falling below B_{lim}. The ICES approach to determining F_{MSY} has been used by the EU for north east Atlantic and Baltic Sea fish stocks and has been written into the current multiannual plans. IUCN suggests that a criterion A1 can be used instead of A2 when F/F_{MSY} <1 for the greater of one generation or five years (IUCN, 2017, 5.6.2.), which does not account for remaining risk to the biomass even if F/F_{MSY}<1.

6.4 Assessing extinction risks

Previous ICES advice (both to Norway and HELCOM) explored the concept of using criterion E to assess the IUCN status of fish (chapter 5 of this report). ICES (2009) advised that PVA approaches assess relative rather than absolute risk to extinction. It also highlighted that continued recruitment below the re-placement line (at low stock size) implies a high probability of extinction, a key dynamic in a population that fisheries management works to avoid. ICES (2013) advised that criterion E should be used to assess status rather than criterion A; this relates back to the issues of evidence base and ICES interpretation of the precautionary approach. Screening approaches are more likely to increase the rate of false positives (prediction of extinction which does not occur) and false negatives (the occurrence of unpredicted extinction).

According to the IUCN guidelines, a quantitative analysis is defined as any form of analysis that estimates the extinction probability of a taxon based on known life history, habitat requirements, threats and any specified management options. The model should include demographic structure, incorporate uncertainty (usually stochastic) and be realistic in terms of the ecology of the organism (IUCN 2017, section 9.2). It is difficult to see how this is very different in concept from management strategy evaluations currently used in fisheries, where risk of management options are explored and trade-offs evaluated.

The guidance (IUCN 2017, sections 9.3 and 9.4) recommends that the data types used to populate the analysis include spatial distribution of suitable habitat, habitat relationships, abundance estimates from surveys, vital rates (fecundity and survival), temporal variation in parameters. The analysis should account for density dependence in demographic parameters, temporal variability of the environment and account for spatial components. Other than the considerations about spatial extent of habitat, all of these issues are built into most standard stock assessments and MSEs carried out by ICES. The stock to recruit relationships are built in, many processes are assumed to vary with stochasticity, and idenitifed, key future threats are explored.

6.5 Updating of assessments and changing of category

Most stock assessments leading to fisheries advice occur within a pre-agreed timeframe and cycle. Many are annual, and almost all within a regular 5 year time frame. IUCN has committed to a 10 year cycle but this requires resourcing to maintain. This resourcing is a challenge to the IUCN system (Rondinini, *et al*., 2014) and the cycle is not automatic.

By interpreting the precautionary approach in a certain manner, there is a bias towards keeping species in more threatened categories. After an update, if a species changes category, that category must occur for at least 5 years before being moved from a higher to lower threat, whereas the species must be re-categorised immediately if moving from lower to higher threat (IUCN 2017, section 2.2.1). Thus while there appears to be no bias towards the threatened categories from the analytical approaches used by IUCN, there is a small bias cause by the process of moving between categories.

Fisheries management methods build precaution into the assessments and advice by incorporating buffers that account for uncertainty and risk.

6.6 Generation time

Listing a species using the IUCN criterion A requires calculation of the rate of decline in population size over the past 10 years or three generation lengths, whichever is longer. Thus accurate determination of generation length can be integral to whether a species will be given a threatened IUCN listing. The correct generation length to use has been disputed.

In exploited fish species it is common that a decline in population size may have occurred in the past (prior to 3 generations lengths ago) but has become stable due to management.

If a species is managed sustainably, the generation length is likely to stabilize at a shorter length than for the unexploited population. In this case it may be beneficial for management at MSY that the generation length does not revert to pre-exploitation lengths, which would imply lower productivity. In turn, this can lead to an unrealistic rate of decline that triggers the IUCN criterion A erroneously. However, the IUCN guidelines state that “where generation length varies under threat, such as the exploitation of fishes, the more natural, i.e. pre-disturbance, generation length should be used” (IUCN, 2017).

The rationale for this is that using the current, shorter generation length, would lead to a lower threat category (because a shorter period is used to calculate the reduction). Keeping the pre-exploitation generation length is more precautionary in this instance, but may not best reflect fish species being sustainably managed (that is, managed at MSY).

6.7 Conclusion on differences in approaches

As concluded by various studies, the IUCN and fisheries management approaches are two complementary approaches that offer differing insights and levels of information, and should not necessarily be taken as directly comparable (see section 3). The biggest differences are

- In the different uses of the evidence base (linked to different interpretations of the application of the precautionary principle),
- the use of reference points for assessing the state of the population (fisheries management) compared to the use of recent trends in population abundance (IUCN),
- the acceptance of the use of projections in fisheries management to assess the risk of proposed management actions (i.e. the applicability of fisheries management strategy evaluation (MSE) as a tool to assess criterion E, extinction risk).

The warning to avoid an evidentiary attitude (i.e. the best estimate) in the IUCN guidelines (2017), suggests a different approach to risk and application of the precautionary approach. In fisheries science, the method most suitable to the assessment of a stock is chosen based on the characteristics and quality of the data. ICES in 2009, suggested that false positives, or false negatives (or false alarms or misses as labelled in chapter 7 of this report and Dulvy *et al.*, 2005) were more likely when assessing using methods like criterion A compared to a stock assessment. Precautionary is built into fisheries management approaches by incorporating uncertainty and adding buffers to reference points that reflect the acceptable risk and testing future scenarios against those reference points. More recently, risk to a population is assessed through stochastic simulations of future population dynamics (MSE). This will be further explored in chapters 7 and 8.

Fisheries management under the CFP generally does not rely just on the recent trend of a stock, but usually assesses the current estimates of F and SSB against reference points, or constructs a tested harvest control rule, that keeps the stock at an acceptable abundance. In contrast, as highlighted before, a stock could be below biomass reference points for a long time and be classified as least concern by IUCN. Such a stock would usually be subject to a rebuilding plan in the context of fisheries management as fails a key management objective. To simplify the message – IUCN is concerned about rates of decline, fisheries management about position relative to targets and limits related to the estimated status and productivity of a stock. This will be further explored in chapter 8.

Criterion E is rarely used by IUCN to assess extinction risk (Dulvy *et al.*, 2005, Kent Carpenter, IUCN, pers comm). In previous advice (2009, 2013), ICES suggested that there was enough information from an age based stock assessment to test extinction risk. A stock assessment is typically conducted for the whole stock and is devoid of spatial information, but many stocks across a geographic area may be assessed (e.g. ICES assesses 15 cod stocks across the North East Atlantic), thus providing a spatial set of units for overall assessment. IUCN have yet to address in their guidance whether a stock assessment with a stochastic management strategy evaluation into the future, can form the basis of an assessment of extinction risk. This will be further explored in chapter 9.

As discussed in earlier sections, there are also minor differences associated with the specific nature of fisheries science, such as stocks being assessed rather than species, the frequency of data collection and the impact of fishing on generation time.

7 Comparison (IUCN and CFP) of the classification of EU stocks

A comparison of the IUCN and fisheries management assessments for European commercial fish (excluding Mediterranean fish) was carried out following the general methods of Fernandes *et al*. (2017) (see Annex 1).

7.1 Background

Assessments of the status of fish in the North East Atlantic as carried out by ICES, ICCAT and IUCN were compared. The Red List categories from the most recent IUCN listing of species of European marine fish (2015) were compared with the most recent 2017 ICES and ICCAT estimates of status B_{pa} and F_{MSY} (or proxies).

The total dataset consisted of 268 stocks of 87 species. B_{pa} or proxy reference points were available for 98 stocks, and F_{MSY} or proxy reference points were available for 108 stocks. The IUCN threat category (threatened or not) was then compared with the two ratios F/F_{MSY} and B/B_{pa}. The focus here was on the comparison of the two methods, not on how well the method reflected the state of the "true" population. Four possible outcomes of the analysis were defined with respect to the IUCN definition of the threat of extinction (similar to Dulvy *et al*., 2005):

- **True positive**. The two systems concur: a stock is not sustainably fished/ below biomass reference points and the threat criteria are met for the species
- **True negative.** The two systems concur: a stock is sustainably fished/ above biomass reference points and the threat criteria are not met for the species
- **Miss.** A stock is exploited unsustainably/ below biomass reference points but the threat criteria are not met for the species. This was termed a false positive by ICES (2009).
- **False alarm**. A stock is exploited sustainably/ above biomass reference points but the threat criteria are met for the species. This was termed a false negative by ICES (2009).

7.2 Methods and Data

Data from the ICES stock assessment database was downloaded for the past 5 years to ensure all stocks currently assessed by ICES were included. If B_{pa} reference points were not available, MSY or management plan reference points were used. If F_{MSY} reference points were not available, PA or management plan reference points were used. F reference points are not available for three Icelandic stocks that use harvest rates rather than F_{MSY}. These were converted to F according to F=-ln(1-(HRref point)).

The ICAAT assessments cover 16 stocks. Reference points for these stocks were not always available but the ratios of B/B_{MSY} and F/F_{MSY} were. As a range of values of F/F_{MSY} and B/B_{MSY} are reported in ICAAT assessment, averages were taken.

IUCN assessments do not currently assess Norway lobster (*Nephrops norvegicus*), and as a result, it could not be included in this study. Yellowfin tuna is likewise not assessed in the European Red List. Both Norway lobster and yellowfin tuna have a global classification only.

7.3 Results

When comparing biomass reference points (B_{pa} or proxy) against IUCN threatened criteria, 72% of stocks were similarly assigned. Three stocks produced a false alarms:

Golden redfish (*Sebastes norvegicus*) in subareas 5,6,12, and 14; Blue Ling (*Molva dypterygia*) in subareas 6a, 7 and Division 5b: and Turbot (*Scophthalmus maximus*) in Subarea 4 (Figure 7.1).

20 stocks were classed as a status miss, that is, IUCN did not assess them as threatened but stock assessment indicated biomasses lower than B*pa*.

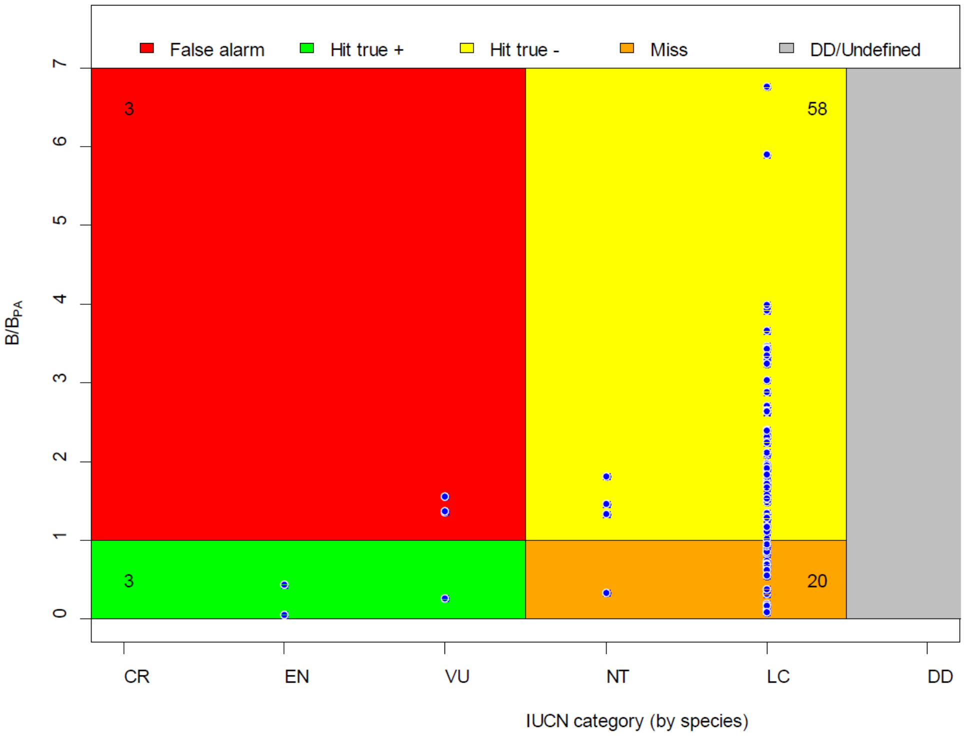


Figure 7.1. Performance of the IUCN Red List in relation to stock status as defined by the estimate of biomass relative to its precautionary reference point (B/BPA, on y axis). Each point is a stock, with the species classified according to the threat criteria of the IUCN Red List (x axis). Red List categories are Critically Endangered (CR), Endangered (EN), Vulnerable (VU), Near Threatened (NT), Least Concern (LC) and Data Deficient (DD). Shading indicates: hits, where the two systems concur, either because the spawning stock biomass is below the reference point and the threat criteria are met (true positive, in green), or because a stock is above the reference point and the threat criteria are not met (true negative, in yellow); misses, in orange, where a stock is at low biomass but does not meet the threat criteria; and false alarms, in red, where the stock is at high biomass but the threat criteria are met. Numbers in each quadrant refer to the number of stocks.

When comparing fishing mortality reference points (F_{MSY} or proxy) against IUCN threatened criteria, 55% of stocks were similarly assigned. Four stocks produced a false alarms: Blue ling (*Molva dypterygia*) in subareas 6a, 7, and Division 5.b; Porbeagle (*Lamna nausus*) in the Northeast Atlantic; Porbeagle (*Lamna nasus*) in the Northwest Atlantic; and Turbot (*Scopthalmus maximus*) in Subarea 4 (Figure 7.2).

33 stocks were classed as a fishing mortality miss (stock assessments indicated a higher F than F_{MSY} but IUCN assessment did not class them as threatened).

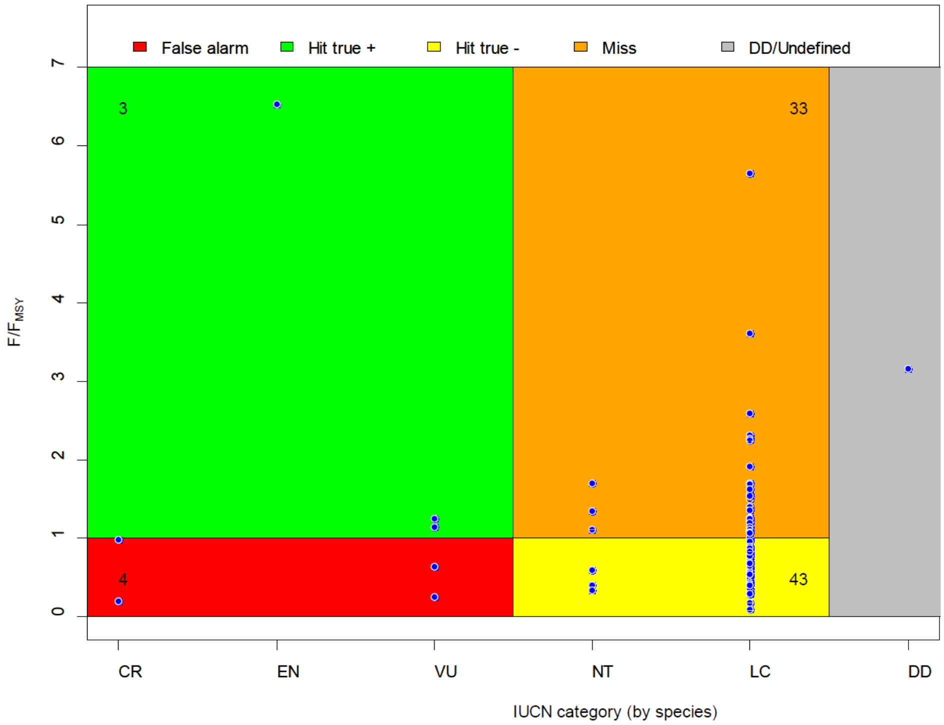


Figure 7.2. Performance of the IUCN Red List in relation to fishing mortality (F) as defined by the estimate of F relative to its MSY reference point (F/FMSY, on y axis). Each point is a stock, with the species classified according to the threat criteria of the IUCN Red List (x axis). Red List categories are Critically Endangered (CR), Endangered (EN), Vulnerable (VU), Near Threatened (NT), Least Concern (LC) and Data Deficient (DD). Shading indicates: hits, where the two systems concur, either because the fishing mortality is above the reference point and the threat criteria are met (true positive, in green), or because the fishing mortality is below the reference point and the threat criteria are not met (true negative, in yellow); misses, in orange, where the fishing mortality is high but does not meet the threat criteria; and false alarms, in red, where the fishing mortality is low relative to the reference point, but the threat criteria are met. Numbers in each quadrant refer to the number of stocks.

7.4 Conclusions

When comparing biomass reference points with status of IUCN threat, the majority of stocks were similarly assigned (72%). However this dropped when comparing fishing mortality reference points to threat status to 55% similarly assigned. So the assessments are more similar when they are based on the same underlying data (e.g. biomass). Both approaches detect changes in biomass to alert concern. However it appears that the IUCN assessment has a reduced ability, in this example, to alert to over-exploitation ($F > F_{MSY}$, as suggested by Punt, 2000).

The number of misses (23% for biomass and 40% for fishing mortality), when the IUCN assessment is compared to the fisheries management reference points suggests that the IUCN approach is less precautionary than the fisheries management reference points. So no bias by IUCN methods towards exaggerating the threat status has been found in this example.

It is necessary to make distinctions which reduce comparability between the IUCN assessments and the stock assessments conducted by ICAAT and ICES. These are:

1. Difference in objectives. IUCN is concerned with extinction risk, whereas fisheries assessments are concerned with sustainable exploitation.

2. IUCN works at the species level whereas ICAAT and ICES work on a stock level.
3. Fisheries assessments deal with a smaller number of commercially exploited species, compared to the IUCN list that covers all species.
4. Fisheries assessments are conducted more frequently than IUCN assessments.

These factors can lead to discrepancies between the two systems.

In some cases, the false alarms can be attributed to the difference in timing between the stock assessments and the IUCN assessment. The most recent IUCN assessment was carried out in 2015, compared to the stock assessments up to 2017.

The differentiation between stock and species impacts the assessments. A particular example highlighted in the report is cod (*Gadus morhua*), classed as Least Concern (LC) by the IUCN system. Whilst some stocks of cod in the Northeast Atlantic are very large (i.e. NE Arctic cod), others have been in decline for many years (e.g. west of Scotland cod). The latter, if assessed by IUCN on a stock basis rather than as a species, would result in a classification as Critically Endangered (CR) under criterion A (a reduction in population size based on an estimated population size reduction of ≥90% over the last 10 years or three generations, whichever is the longer).

In relation to commercially exploited fish, the two systems can be seen as "complementary graduated indicators of status", with the fisheries assessments representing the first level of concern and status in relation to sustainable exploitation. As key objectives of the CFP include that fishing is environmentally sustainable in the long-term, with the minimisation of negative impacts, it may be beneficial to append the IUCN assessment to the stock assessment as a matter of course for all stocks. Potentially the Red List could be updated more regularly using the stock assessment, however, the issue of stocks vs species would remain.

8 A comparison between IUCN categories of conservation status and fisheries reference points

8.1 Background

A comparison of the IUCN categories of conservation status and fisheries management assessments was carried out using a Management Strategy Evaluation framework. As suggested in chapter 7 (and annex 1), the extinction risk to a stock is likely influenced by the life history characteristics of the species. So when comparing between IUCN categories of conservation status and fisheries reference points, a range of life history characteristics should be explored.

A series of stock projections were conducted for four stocks: western Atlantic Bluefin tuna (*Thunnus thynnus*), blue shark (*Prionace glauca*), cod in the Baltic Sea (*Gadus morhua*), and Blackspot seabream (*Pagellus bogaraveo*) (for the full report see Annex 2). This method simulates a "true" population, about which the stock dynamics are known. It is then possible to assess how this simulated stock responds to management using an ICES-style harvest control rule.

A simulated pseudo-assessment was used in conjunction with an ICES-style harvest control rule to project the four stocks forward 100 years, subject to process error and estimation uncertainty. Simulations were compared using the stock assessment estimates of spawning stock biomass relative to biomass reference points. The time series were also assessed for population decline using IUCN criterion A. The distribution of stock outcomes from 1000 simulations for each species was also used to compute probabilities of extinction for assessment under IUCN criterion E.

8.2 Methods and Data

A Management Strategy Evaluation (MSE) was performed for the four stocks. In each case, an operating model (OM) was developed from current information on stock characteristics and conditioned on available information and catch history to represent a simulated "true" population (see annex 2 for the OM specifications). Populations were assumed to be unexploited in the initial years of the simulation. Each OM was projected forward 100 years, with catches informed by a pseudo-stock assessment and a harvest control role designed to replicate that frequently used by ICES.

A TAC for each year was then calculated by comparison of the SSB to an estimate of MSY $B_{trigger}$, to give an estimate of fishing mortality rate. The control rule used was:

$$F_y = \begin{array}{ll} F_{MSY} & if\ \widehat{SSB_y} \geq MSY\ B_{trigger} \\ F_{MSY}\widehat{SSB_y}/MSY\ B_{trigger} & if\ \widehat{SSB_y} < MSY\ B_{trigger} \end{array}$$

These estimates of F_{MSY} and MSY $B_{trigger}$ were obtained from the OM in the initial year of the projection period, by assuming MSY $B_{trigger}$ = 0.5BMSY, with deterministic estimates of BMSY and FMSY obtained given simple yield-per-recruit analysis. This approach is not identical to that used by ICES (which is stochastic) but was more computationally feasible.

After the value for F_y is calculated, this is then used with the current estimate of spawning biomass to calculate the total catch over ages that would be expected given this value of fishing mortality to define the TAC:

$$TAC_y = \frac{\widehat{SSB_y}}{SSB_y} \sum_{l=1}^{N_L} \sum_{s=1}^{2} w_{s,l} \sum_{a=0}^{A} \Phi_{l,s,a} N_{s,a,y}\, e^{-0.5M_a}(1 - e^{-S_a F_y})$$

where S_a is the selectivity at age a, $w_{s,l}$ is the weight at length bin l for sex s, $\Phi_{l,s,a}$ is the proportion of fish of sex s and age a in length bin l, N and M are the numbers and mortality at age, and N_L is the number of length bins, and A is the maximum age. This approach therefore assumes that the catch is calculated given the correct (OM) age structure, but does account for the effect of assessment error on the TAC given a value of F_y from the control rule. This TAC was fed back into the OM to update the population dynamics for the next year. In total, 1000 simulations were conducted for each species and OM scenario.

In each year of the projection, the estimated SSB was compared to reference points and the magnitude of population decline over the previous 3 generations was calculated (as is required for calculation of IUCN criteria A). Additionally, extinction risk was also calculated from the 1000 simulations (enabling comparison of population simulations under IUCN criterion E).

For each stock, for each assessment, two sets of performance matrices were calculated to show:

- If IUCN criterion A thresholds were met
- If the assessment estimate was below biomass reference points (B_{pa} and B_{lim})

To estimate the magnitude of population size decline needed for the calculation of IUCN criterion A, generation time was computed for the four stocks. This was achieved by comparing assessment estimate of SSB in the current year to that at time Tgen+1.

Comparison to the biomass reference points was achieved by comparing the assessment values of SSB to the estimated reference points.

An additional set of metrics to evaluate the probability of extinction risk from a quantitative analysis were calculated by considering the distribution of outcomes over simulations for each stock scenario. The frequency in which SSB in the terminal year of the projections fell below 10%, 5% and 1% of the unfished SSB was calculated.

8.3 Results

8.3.1 Western Atlantic Bluefin tuna

During the historical period, operating model values for the magnitude of biomass change exceeded IUCN thresholds for criterion A in a large number of cases as the stock is fished down. After implementation of the management period, this number was greatly reduced. At the end of the simulation period, the stock began to oscillate around biomass reference points, which triggered a threatened listing more often than at the beginning of the management period (see figure 8.1b and c).

Whilst IUCN criterion A listing were triggered infrequently at the start of the management period, the stock was below reference points (B_{pa} and B_{lim}) in almost all simulations, due to the low initial level of the stock. The stock remained below reference points in the majority of simulations for the entire projection period, reflecting a misspecification in the deterministic reference points used in this simulation work (Figure 8.1d). As the management period continued the frequency at which the stock was estimated to be below B_{lim} decreased. These results show that status in relation to reference points more accurately reflects the true nature of the depleted stock compared to IUCN category A listing. The analyses also showed that whether a stock was given an IUCN listing did not correspond to whether the stock was also below biomass reference points, highlighting the lack or correspondence between IUCN listing and status in relation to reference points for this stock (see the confusion matrices in Figure 6 in Annex 2).

Only 1 out of 1000 simulations showed the stock falling to below 10% of unfished spawning biomass during the simulations, showing that even this mis-specified control rule which maintained the stock at low levels would not lead to extinction. Thus if the stock was assessed using IUCN criterion E it would not trigger a threatened listing

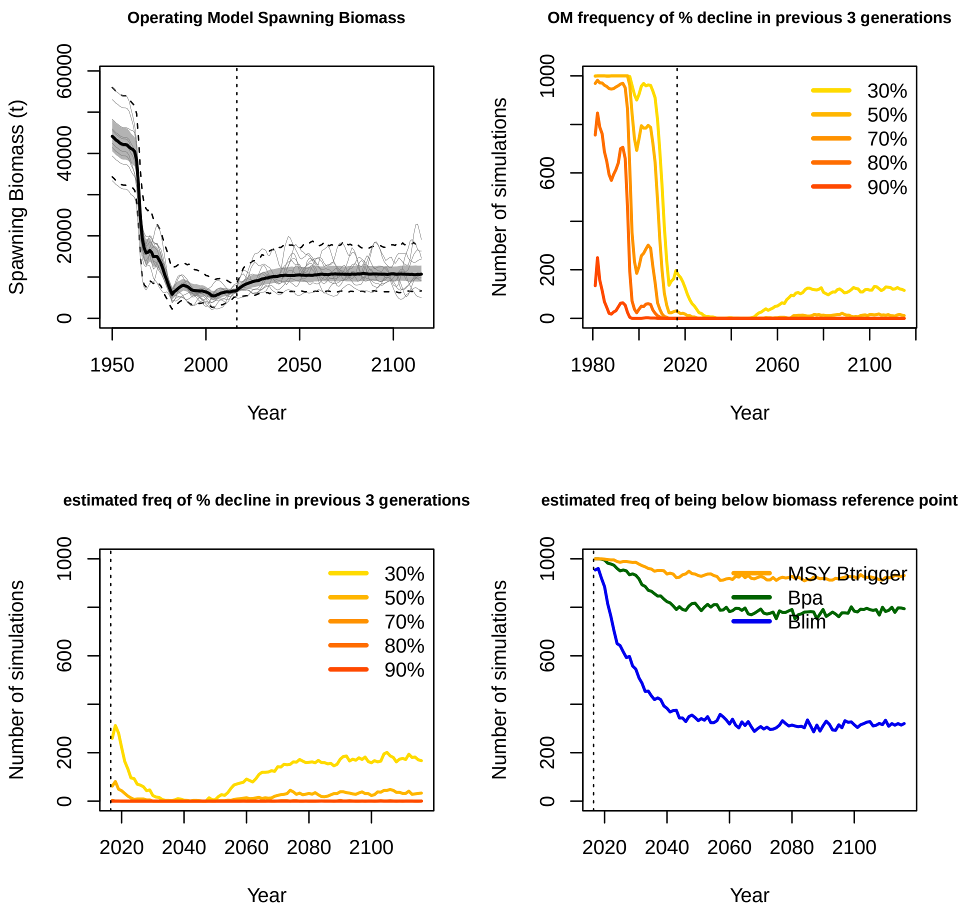


Figure 8.1 Time series of outputs for western Atlantic Bluefin tuna simulations. (a) summary of spawning biomass trajectory, with median (dark line), interquartile range (grey shading), and 95% interval (dashed lines). Example individual simulation trajectories shown in grey lines. (b) Number of simulations in which change in operating model spawning biomass over three previous generations exceeds thresholds for IUCN categories, (c) Number of simulations in which the estimated change in \spawning biomass over three previous generations exceeds thresholds for IUCN categories, (d) number of simulations in which spawning biomass is estimated to be below ICES biomass reference points.

8.3.2 Blue Shark

Two sets of simulations were run for blue shark: the first assuming the stock was at 25% unfished biomass at the start of the projection period and the second with the stock at 40% unfished biomass.

Throughout the historical period, IUCN threat listings of Vulnerable and Endangered were triggered during most simulations when the stock started at 40% and 25% unfished biomass, and this continued to a lesser degree in the management period (Figure 8.2a and b show results for when the stock started at 40% unfished biomass). The stock was estimated to be below B_{pa} approximately 20% of the time during the projection period when the stock started at 40% unfished biomass and very rarely below B_{lim} (Figure 8.2).

There was more correspondence between IUCN listings and assessment of status relative to reference points for blue shark than Bluefin tuna. In 36% of cases, the stock was estimated to be below B_{pa} and to have triggered the IUCN threat listing using criterion A, and in 65% of cases when compared to B_{lim}. There were very few instances where the >=50% decline threshold was exceeded and the stock was not assessed as being below B_{pa} or B_{lim} (see Figure X in Annex 2).

When the stock was in a better state at the beginning of the projection period (40% unfished spawning biomass) there was more agreement between IUCN threat listing and spawning biomass relative to reference points.

Despite this being a data-limited stock, and therefore having an increased uncertainty in assessment, there were zero instances of simulated stocks falling below 10% unfished spawning biomass (that would trigger IUCN criterion E).

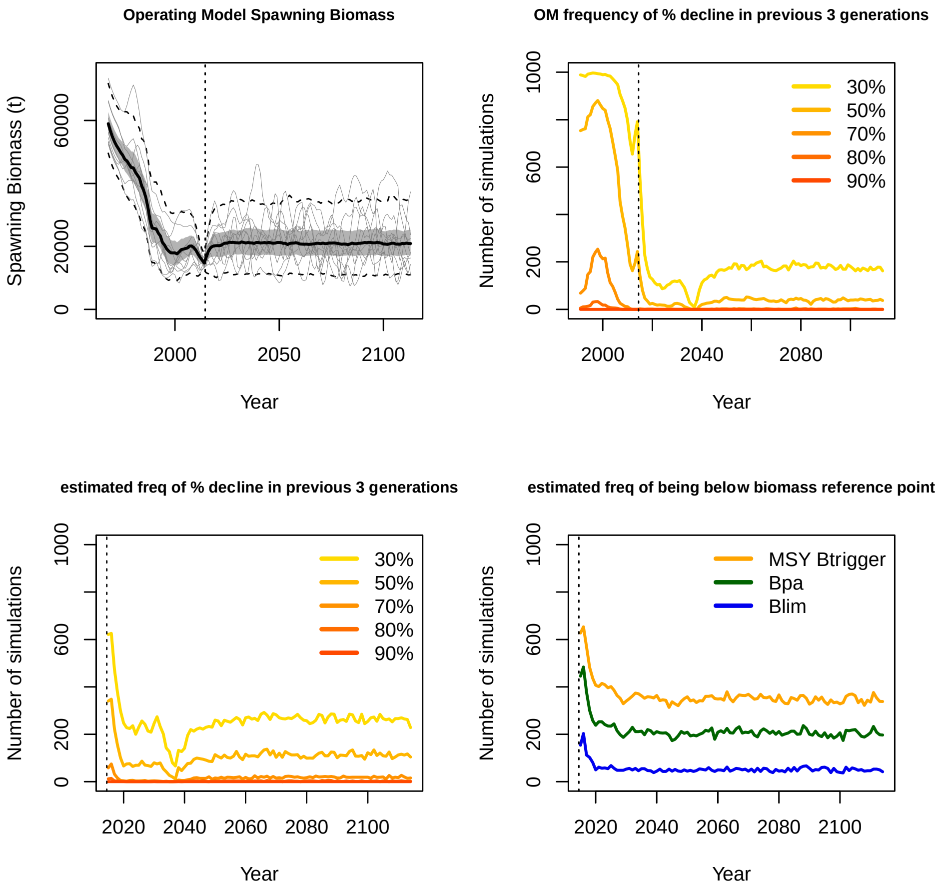


Figure 8.2 Time series of outputs for blue shark simulations assuming that the stock is at 40% unfished biomass at the start of the projection period. (a) Summary of spawning biomass trajectory, with median (dark line), interquartile range (grey shading), and 95% interval (dashed lines). Example individual simulation trajectories shown in grey lines. (b) Number of simulations in which change in operating model spawning biomass over three previous generations exceeds thresholds for IUCN categories, (c) Number of simulations in which the estimated change in \spawning biomass over three previous generations exceeds thresholds for IUCN categories, (d) number of simulations in which spawning biomass is estimated to be below ICES biomass reference points.

8.3.3 Baltic cod

Scenarios were run for both the eastern and western Baltic cod stocks, assuming complete independence between stock dynamics and fisheries. Both were assumed to be at 12.5% unfished spawning biomass at the beginning if the management period. Biological parameters for both stock were assumed to be identical. To combine the stocks, the assessment estimates of biomass from the separate analyses were summed over both stocks to obtain a time-series of biomass estimates for both stocks combined, as per IUCN guidelines (2017).

As for the other stocks, Baltic cod frequently triggered the IUCN threat categorisation in the historical period (Figure 8.3b and c). After the implementation of control rules during the management period IUCN listing only occurred in 20% of cases. However, the stock continued to be estimated to be below $MSYBt_{rigger}$ in approximately 50% of cases (Figure 8.3d). As for bluefin tuna, there was a lack of correspondence between the stock being below biomass reference points and receiving an IUCN threat listing.

Confirming this lack of correspondence, in only 12% of simulated assessments were stocks assessed to be below B_{pa} and showed a >=50% biomass decline (that is, IUCN and assessment results were in agreement that the stock was threatened).

When IUCN biomass decline statistics were calculated from summing eastern and western assessment estimates, the number of simulations in which listing categories were triggered increased only slightly. Instances of exceeding the >=50% threshold were only triggered when both stocks were assessed to be below the biomass reference points. If only one stock is assessed to be below the biomass reference points, then the number of IUCN listings across simulations remains very low or zero. Effectively the larger (eastern) stock masked the smaller (western) stock, showing this method used by IUCN for fisheries stocks may not be appropriate when the smaller stock is not doing as well as the larger stock.

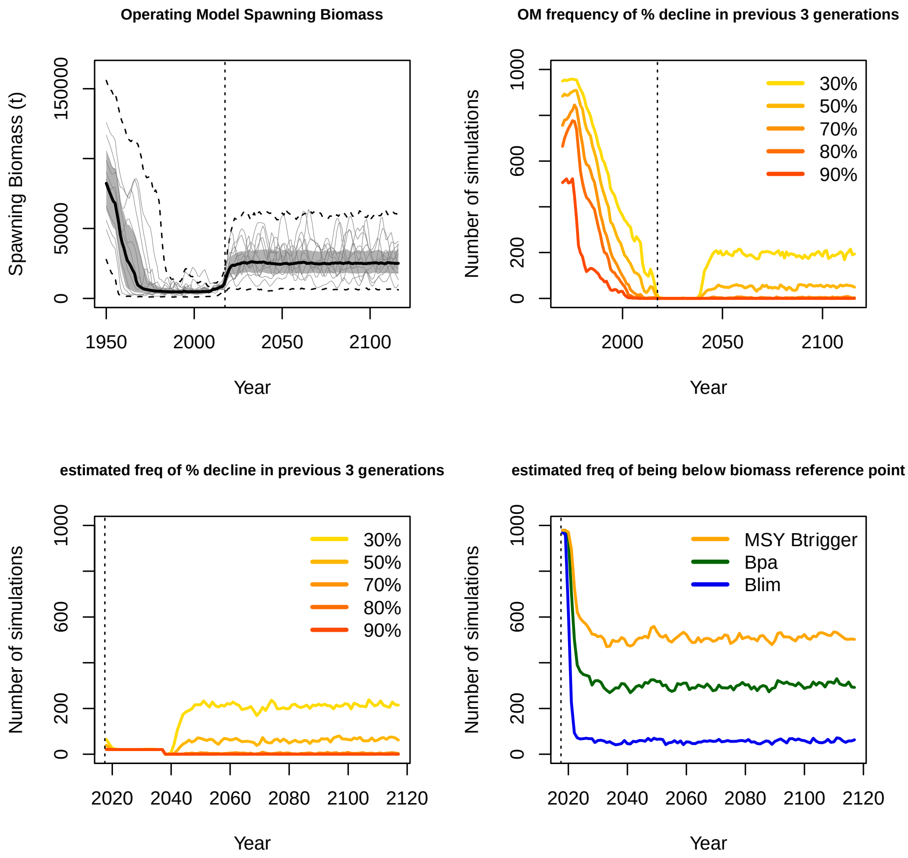


Figure 8.3: Time series of outputs for western Baltic cod simulations. (a) summary of spawning biomass trajectory, with median (dark line), interquartile range (grey shading), and 95% interval (dashed lines). Example individual simulation trajectories shown in grey lines. (b) Number of simulations in which change in operating model spawning biomass over three previous generations exceeds thresholds for IUCN categories, (c) Number of simulations in which the estimated change in \spawning biomass over three previous generations exceeds thresholds for IUCN categories, (d) number of simulations in which spawning biomass is estimated to be below ICES biomass reference points.

8.3.4 Blackspot seabream

Blackspot seabream was treated as a data-limited stock in this analyses, where the pseudo-assessment was assumed to have a CV of 40% and to be heavily depleted at the beginning of the management period (10% unfished spawning biomass).

As for the other stocks, blackspot seabream frequently triggered the IUCN threat categorisation in the historical period (Figure 8.4b and c). The frequency at which the stock was assessed to be below biomass reference points dropped to very low levels during the management period (Figure 8.4d).

The probability of falling below 10% spawning biomass was just under 5% in the early years of the management period reflecting the low initial stock size. This is higher than for the other stocks, reflecting the data-limited nature of the stock and the increased uncertainty.

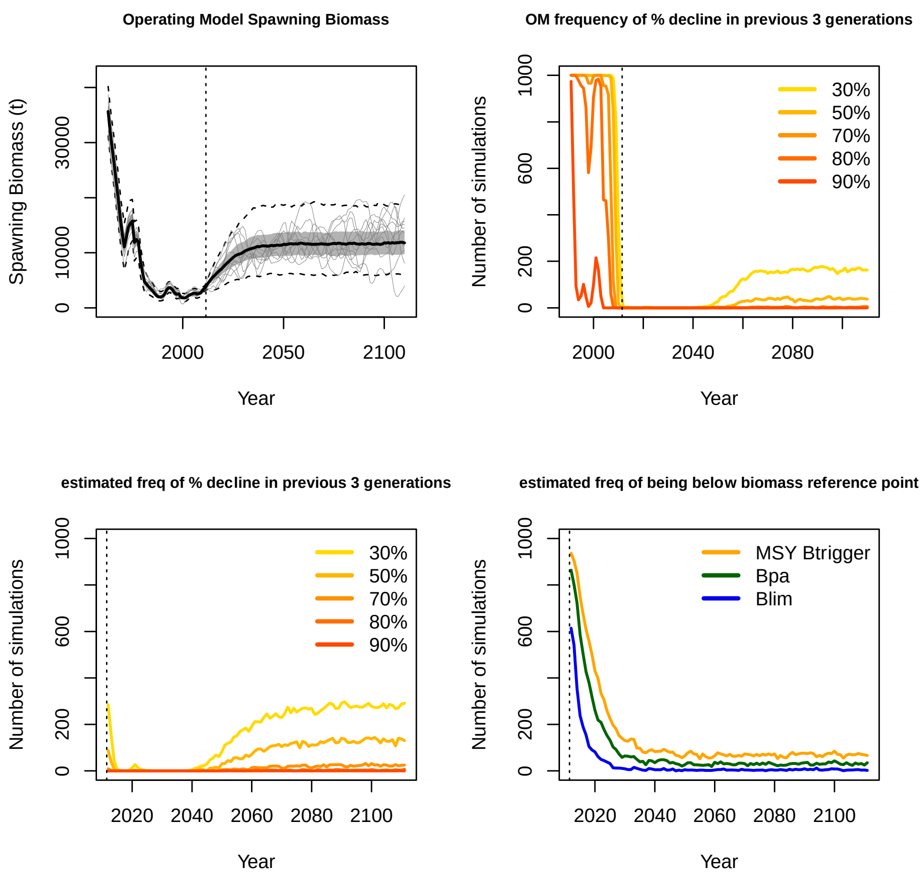


Figure 1: Time series of outputs for blackspot seabream simulations. (a) summary of spawning biomass trajectory, with median (dark line), interquartile range (grey shading), and 95% interval (dashed lines). Example individual simulation trajectories shown in grey lines. (b) Number of simulations in which change in operating model spawning biomass over three previous generations exceeds thresholds for IUCN categories, (c) Number of simulations in which the estimated change in \spawning biomass over three previous generations exceeds thresholds for IUCN categories, (d) number of simulations in which spawning biomass is estimated to be below ICES biomass reference points.

8.4 Assumptions

As with all MSE approaches, assumptions were made during the process. There was an assumption made in the initial years of the simulations that populations were unexploited, and with equilibrium age structure. This is likely not to be the case for stocks that have been exploited for a long time, such as Baltic cod. As a result, it is likely that listings under IUCN criterion A during the historical period are more likely in these analyses than would have been the case if the true underlying stock dynamics were better known. The fisheries for each stock were also modelled as single stocks with single fisheries, which is unlikely to be the case. By not taking this into account it is likely that there is increased bias and imprecision in estimates of stock size and reference point determination. The pseudo-assessment approach, whilst computationally more simple, does not take into account the full structure and characteristics of a stock assessment. Nor do the approximations to the ICES-style harvest control rule and reference point determination. The result of these differences is that the reference points in the analyses were likely to be less precautionary than ICES reference points (meaning a stock would be assessed in the simulations presented here as being below such reference points less often). However, studies (such as Punt *et al*. 2008) have shown that the approach used here is suitable for approximating the behaviour of stock assessment processes.

8.5 Conclusions

A key management objective of the CFP is the maintenance of populations of harvested species above levels that can produce the maximum sustainable yield. This requires an assessment of the state of the stock against biomass reference points. The MSE simulation work conducted here, demonstrated the utility of assessing stocks against reference points reflecting stock status and linking those to a harvest control rule, rather than assessing only stock trend. The simulations were based on populations with a range of life history characteristics and two assumptions about certainty of the data. A trigger of an IUCN threatened status was infrequently in the management period of the simulations, despite stocks still being below biomass reference points, due to a lack of further declines in population size. Thus, the assessment approach was seen to be more precautionary and alert to crossing limit thresholds more regularly than the IUCN threat categorisation.

As expected, the IUCN threat listing was triggered frequently when assessing *historical* trends in biomass for these stocks due to a declining population size at the beginning of exploitation. Populations that began management projection periods at a higher stock level showed greater correspondence between IUCN category listings and stock status relative to reference points than stocks at a low level. When the harvest control rule estimated stocks to be above biomass reference points, the IUCN assessment also tended to give positive results. Towards the end of the 100 year stock projections, listing of Vulnerable again became more frequent, as populations were on average at higher levels, and moving around biomass targets.

Analyses for Baltic cod show that there is a problem summing over multiple populations. The IUCN criteria was not triggered despite one of the two stocks being below biomass reference points.

The work also demonstrated that MSE can be a more comprehensive tool for assessing uncertainty associated with management of social-ecological systems than the traditional Population Viability Analysis (PVA), suggesting MSE is a useful means for assessing fish populations under IUCN criterion E. The analyses show that the harvest control rules applied here were successful at preventing stock crash and maintaining stocks above critical biomass limits, even in data-poor cases. Simulated stocks very rarely, or never, fell below the 10% spawning biomass threshold. Using this analysis to assess a stock against criterion E would thus not lead to any of these 4 stock being listed, whereas assessment using criterion A (population decline) would have done so in many cases during the historical period when the stock was being fished down.

However, when stocks such as Bluefin tuna were assessed to be below biomass reference points in the management period in nearly all cases, a change in IUCN status would still not have been triggered. Thus, whilst criterion E is shown to be a useful means to assess extinction risk, it is not useful for the management objective of maintaining a biomass above a certain limit, unless it is associated with other criteria for assessment.

9 Summary of conclusions in preparation to answer the request

9.1 Use of data and information in IUCN assessment process for marine species compared to processes of relevant fisheries organisations

When assessing commercial stocks for fisheries management, many time series of data from different sources are used to estimate stock status. It is useful to view a stock assessment as an integrator of the available empirical data (catch monitoring, biological sampling, surveys, tagging or catches per fleet). The time series of biomass could be considered as an integrated index of multiple data streams. IUCN guidance recommends assessing a population against all the criteria, but in practice commercial fish stocks are assessed only using criterion A (Dulvy *et al*., 2005), and very rarely criterion B (geographic range for endemic stocks, e.g. Norwegian fjord populations and Australian hand fish). The data used for an IUCN assessment of a commercially exploited species tends to be the output of a stock assessment, usually the biomass trends. Both fisheries and IUCN consider mature individuals.

It could be argued that currently, by using criterion A and not E, the IUCN approach discards information that helps estimate recent trends and risk to extinction (such as age or length structure, recruitment of juveniles and exploitation rate). What ICES in 2009 called "screening methods", which includes the IUCN approach, are more prone to false alarms than the best estimates from fisheries management approaches. But IUCN say this is an important aspect of their interpretation of the precautionary approach. This would only be the case if there were more false alarms than misses, which does not appear to be the case currently in Europe where misses have been more common than false alarms (see chapter 7) and continuous long periods of low biomass will not alert the IUCN categories (see chapter 8).

The data requirements for criterion A are similar to the data available for many data-deficient fish stocks; with an appropriate index of abundance over a time series being used to assess trends. Much effort has recently been expended exploring methods for assessing data-deficient stocks (e.g. ICES WKLIFE workshops, Anderson *et al*., 2017 etc.). The performance of these methods has not been compared to criterion A approaches. The IUCN guidance recommends survey indices over commercial catch per unit effort (CPUE), in recognition of some problems that can impact CPUE time series.

9.2 Criteria used by IUCN and their suitability for assessing marine fish species conservation status.

ICES has previously assessed the suitability of the IUCN criteria (2009) and concluded that for fish with a stock assessment, criterion E was most appropriate. However, IUCN warn against an "evidentiary attitude" and recommend that as many criteria as possible be used to assess the status of the species. If both criteria A and E are used and criterion A results in a higher threat category, then the assessment using criterion A will be adopted, despite the likely evidence base being weaker. Simulations suggest that criterion A is likely to be trigged on some occasions when populations have recovered and are moving around a biomass reference point.

The studies referred to in this report, Rice and Legace (2007) and Davies and Baum (2012) illustrate that the IUCN analytical approach is not biased towards exaggerating the threat status. However the process of moving between categories may add a bias (see section 6.5).

Overall, the IUCN criteria do broadly reflect the assessments of stock status derived for biomass trends from fisheries management approaches.

9.3 IUCN categories of conservation status compared to fisheries reference points for defining sustainability at both the species and stock level

The IUCN and the fisheries approaches are methods to address difference objectives. This results in the IUCN assessments being concerned about rates of decline, and the fisheries assessments about status relative to biomass targets and/or limits.

Simulations suggest that the IUCN threatened status will be infrequently triggered in a stock that is stable below biomass reference points, due to a lack of further decline in population size. That is, when a stock has been depleted and estimated to be below biomass reference points for a long time (greater than 3 generations or 10 years), it will not be categorised as threatened using the IUCN guidelines. This is a result of the differing objectives of each system. In the fisheries sustainability context, it could be said that fisheries assessment approach is more precautionary and accounts better for the risk to a stock being outside safe biological limits than the IUCN threat categorisation.

The simulations suggest that populations at a higher stock level at the beginning of the management projection period were likely to show greater correspondence between IUCN category listings and stock status relative to reference points than stocks at a low level. When the fisheries management approach estimated stocks to be above biomass reference points, the IUCN assessment also tended to give similar positive status. Towards the end of the 100 year stock projections, listing of Vulnerable became more frequent, as populations were on average at higher levels, and moving around biomass reference points.

By testing for extinction risk (with 1000 simulations over 100 years), the simulations suggest that the harvest control rule prevents stock crash and maintains stocks above critical extinction biomass limits, even in data-poor cases. Simulated stocks very rarely, or never, fell below the 10% spawning biomass threshold. Using this analysis to assess a stock against IUCN criterion E would thus not lead to any of these 4 stocks being listed. Whereas an assessment using criterion A (population decline) would have led to a listing for a few simulated populations. Taking into account the warning about evidentiary attitude in the IUCN guidelines, these few simulated populations would be listed, although there was no risk to the populations becoming extinct, under assumed conditions. It is important to emphasise that in fisheries science, the evidence is regularly updated through a monitor and assess approach, so the underlying assumptions about environmental conditions and population dynamics are regularly tested and updated.

The work also demonstrated that MSE can be a more comprehensive tool for assessing uncertainty associated with management of social-ecological systems than the traditional Population Viability Analysis (PVA), suggesting MSE is a useful means for assessing fish populations under IUCN criterion E.

9.4 Recent IUCN listings for key EU stocks with corresponding information from stock assessments

The comparison of assessments of status from IUCN and fisheries management using biomass, shows a general coincidence suggesting that the two approaches have similarly assigned the stock status with regards to biomass, in Europe. Although almost one in four stocks were below B_{pa} (or proxy) and assigned to non-threatened by the IUCN approach. Davies and Baum (2012) found a miss-classification with 13.5% of Red Listed marine fishes classed as threatened, while 40% of populations with a stock assessment were classed as being below biomass limit reference points.

When using fishing mortality reference points, it appears that the IUCN assessment has a reduced ability to alert to over exploitation. This highlights that as a conservation tool, the IUCN approach is not designed to advise on maximising yield. Punt (2000) suggestion that the IUCN approach may not identify species at risk of extinction needing fishing mortality to be reduced.

The number of misses (23% for biomass and 40% for fishing mortality) suggests that the IUCN approach underestimates risk to a stock and compared to a fisheries management reference points approach it cannot fully aid delivery of CFP objectives. There is no evidence of bias by IUCN analytical methods towards exaggerating the threat status in this example.

9.5 Assess suitability and reliability, and advantages and disadvantages of incorporating IUCN assessments for fisheries management and other relevant decision making processes

All relevant published research has emphasised that IUCN listing and fisheries management approaches have different but complementary objectives. These different objectives result in different interpretation of the precautionary principle, and different attitudes towards data and the knowledge base. By generally using criterion A to assess marine fish, IUCN focuses on the direction and magnitude of the recent trend in biomass (using stock assessment output or survey indices). This operates well to alert management authorities to recent change. It could be considered the equivalent of a data poor warning bell in the context of commercially exploited fish and is a crucial alarm for species that are not-commercially exploited.

In terms of providing information for the maximising yield for fisheries and assessing risk to fish stocks in relation to biomass reference points, the IUCN approach does not perform well, in that is provides no information directly relevant to these purposes. It cannot detect that a stock is overfished if the biomass is very low and stable. It could be argued that when a fish stock is exploited, the IUCN approach can underestimate the risk to a stock being outside safe biological limits.

IUCN has the objective to list all species. For commercially exploited fish, this can be easily achieved by coupling an IUCN assessment to a stock assessment. This would supplement the knowledge base to assessing the CFP objectives of long term sustainability and minimising negative impacts of fishing activities. This would aid Europe in assessing progress towards the Aichi targets and UN SDG 14. If MSE was accepted by IUCN as a valid method for assessing criterion E, then assessing the extinction risk could be easily coupled to the fisheries management process.

Simulations suggest that some data deficient stock assessments perform better than IUCN approaches to alert for extinction risk (see blue shark and black spot bream examples) and when associated with management action, they prevent biomass remaining below reference points.

The example of Norway, shows that the stock assessment and IUCN approach can be built together to provide a robust warning system for marine fish. If used alone, IUCN approaches would not be reliable as an evidence base to deliver the objectives of the CFP, they would need to be supplemented with approaches that consider the status on the biomass relative to reference points and some consideration of the exploitation/by-catch rate.

10 References

Anderson, S.C., Cooper, A.B., Jensen, O.P., Minto, C., Thorson, J.T., Walsh, J.C., Afflerbach, J., Dickey-Collas, M., Kleisner, K.M., Longo, C., Osio, G.C., Ovando, D., Rosenberg, A.A., Selig, E.R. 2017. Improving estimates of population status and trajectory with superensemble models. Fish & Fisheries 18: 732-741 DOI:10.1111/faf.12200

Cadrin, S. X., and Dickey-Collas, M. 2015. Stock assessment methods for sustainable fisheries. ICES Journal of Marine Science, 72: 1–6.

Davies, T. D. and Baum, J. K. 2012. Extinction Risk and Overfishin: Reconciling Conservation and Fisheries Perspectives on the Status of Marine Fishes. Scientific Reports 2:561.

Dulvy, N. K., Jennings, S., Goodwin, N. B., Grant, A. and Reynolds, J. D. 2005. Comparison of threat and exploitation status in North-East Atlantic marine populations. Journal of Applied Ecology 42: 883-891.

EU 2013. Regulation (EU) No 1380/2013 of the European Parliament and of the council of 11 December 2013 on the Common Fisheries Policy, amending Council Regulations (EC) No 1954/2003 and (EC) No 1224/2009 and repealing Council Regulations (EC) No 2371/2002 and (EC) No 639/2004 and Council Decision 2004/585/EC. 40pp.

Fernandes, P. G., Ralph, G. M., Nieto, A., Criado, M. G., Vasilakopoulos, P., Maravelias, C. D., ... Carpenter, K. E. (2017). Coherent assessments of Europe's marine fishes show regional divergence and megafauna loss. Nature Ecology & Evolution, 1, [0170]. DOI: 10.1038/s41559-017-0170

Gullestad, P., Abotnes, A., Bakke, G., Skern-Mauritzen, M., Nedreaas, K., Sovik, G. 2017. Towards ecosystem-based fisheries management in Norway – Practical tools for keeping track of relevant issues and prioriting management efforts. Marine Policy 77: 104-110.

Hilborn, R., and Walters, C. J. 1992. Quantitative Fisheries Stock Assessment: Choice, Dynamics, and Uncertainty. Chapman and Hall. 570 pp.

Hudson, E. 1999

Hudson, E., and Mace, G. 1996. Marine fish and the IUCN red list of threatened animals. Report of the workshop held in collaboration with World Wildlife Fund (WWF) and IUCN at the Zoological Society of London 29 April – 1 May 1996. WWF and IUCN. London.

ICES. 2009a. Report of the ICES Advisory Committee 2009. ICES Advice 2009. Book 1, pp. 147–156.

ICES. 2009b. Report of the Workshop on analytical methods for evaluation of extinction risk of stocks in poor condition (WKPOOR1), ICES Headquarters 18–20 May 2009 (ICES CM 2009/ACOM:29)

ICES. 2009c. Report of the Workshop for the Exploration of the Dynamics of Fish Stocks in Poor Conditions (WKPOOR2), Bergen, Norway 24–27 August 2009 (ICES CM 2009/ACOM:49)

ICES 2013. Special Request February 2013. 8.3.4.1 Review of HELCOM draft Red List assessment of cod (Gadus morhua). XXXX

IUCN 2017. Guidelines for using the IUCN Red List Categories and criteria. Version 13. Prepared by the IUCN standards and petitions subcommittee. Downloadable from http://cmsdocs.s3.amazonaws.com/RedListGuidelines.pdf. 108pp

Keith, D., McCarthy, M., Regan, H., Regan, T., Bowles, C., Drill. C., Craig, C., Pellow, B., Burgman, M. A., Master, L. L., Ruckelshaus, M., Mackenzie, B., Andelman, S. and Wade, P. R. 2004. Protocols for listing threatened species can forecast extinction. Ecology Letters 7: 1101-1108.

McCauley, DJ, Pinsky, ML, Palumbi, SR, Estes, JA, Joyce, FH, & Warner, RR. (2015). Marine defaunation: Animal loss in the global ocean. Science, 347(6219). http://dx.doi.org/10.1126/science.1255641

Musick, J. A. 1999. Criteria to Define Extinction Risk in Marine Fishes. The American Fisheries Society Initiative 24 (12): 6-14.

Powles, H., Bradford, M. J., Bradford, R. G., Doubleday, W. G., Innes, S., Levings, C. D. 2000. Assessing and protecting endangered marine species. ICES Journal of Marine Science 57: 669-676.

Punt, A. 2000. Extinction of marine renewable resources: a demographic analysis. Population Ecology 42: 19-27.

Quinn, T. J., II, and Deriso, R. B. 1999. Quantitative Fish Dynamics. Oxford University Press, New York. 654 pp.

Reiss H, Hoarau G, Dickey-Collas, M, Wolff WG 2009. Genetic population structure of marine fish: mismatch between biological and fisheries management units. Fish and Fisheries 10: 361–395.

Rice, J. C., and Legace, E. 2007. When control rules coolide: a comparison of fisheries management reference points and IUCN criteria for assessing risk of extinction. ICES Journal of Marine Sceince 64: 718-722.

Rodrıguez JP *et al*. 2015 A practical guide to the application of the IUCN Red List of Ecosystems criteria. Phil. Trans. R. Soc. B 370: 20140003. http://dx.doi.org/10.1098/rstb.2014.0003

Romeiras, M.M., Catarino, S., Filipe, A.F., Magalhaes, A.F., Duarte, M.C., Beja, P., 2016. Species Conservation Assessments in Oceanic Islands: the Consequences of Precautionary Versus Evidentiary Attitudes. Conservation Letters, 9(4), 275–280

Rondinini, C., Di Marco, M., Visconti, P., Butchart, S.H.M., Boitani, L., 2014. Update or Outdate: Long-Term Viability of the IUCN Red List. Conservation Letters, 7(2), 126–130

Annex 1 Comparison of European fisheries management status and IUCN status

A COMPARISON OF IUCN AND ICES FISH STOCK STATUS

A REPORT BY DR. PAUL G. FERNANDES

4 JULY 2018

ABERDEEN, SCOTLAND, UK

Declaration: Any opinions expressed in this report are entirely the author's own. They are based on his own experience with stock assessment conducted with the International Council for the Exploration of the Sea (ICES) and the assessment of extinction risk as designated by the International Union for the Conservation of Nature (IUCN).

Front cover: An image of the porbeagle, *Lamna nasus*, © Citron / CC-BY-SA-3.0.

Contents

Executive summary

1) This report aims to compare the assessments of the status of fish in the North East Atlantic as carried out by three organisations: the International Council for the Exploration of the Sea (ICES); the International Commission for the Conservation of Atlantic Tunas (ICCAT); and the International Union for the Conservation of Nature's (IUCN) Red List. ICES and ICCAT assess the status of commercially exploited fish stocks in order exploit them sustainably. IUCN assesses the extinction risk of species, and this can be regionalised, as was the case in the recent assessment which produced the European Red List of Marine Fishes.

2) The specific task was to: "Compare the Red list categories from 2015 IUCN listing of species of European marine fish with the most recent, up to 2017, estimate of status B_{PA} (or proxy) and F_{MSY} (or proxy)) of fish stocks assessed by ICES and ICCAT."

3) Data were taken from the IUCN European Red List of Marine Fishes carried out in 2013 and 2014; the ICES stock assessment database with assessments available from 2012-2018; and ICCAT data from their 2018 summary report. The final dataset consisted of 268 stocks of 87 species. Status (B_{PA} or proxy) reference points were available for 98 stocks. F_{MSY} (or proxy) reference points were available for 108 stocks.

4) The IUCN threat category (threatened or not) was compared with the two ratios F/F_{MSY} and B/B_{PA} reflecting fishing mortality and status as defined by ICES or ICCAT. Four possible outcomes of this analysis were defined with respect to the IUCN definition of the threat of extinction: True positive, where the two systems concur, and where a stock is not sustainable and the threat criteria are met; True negative, where the two system concur, and where a stock is sustainable and the threat criteria are not met; Misses, where a stock is exploited unsustainably but does not meet the threat criteria; False alarms, where the stock is exploited sustainably but the threat criteria are met.

5) In terms of status (B_{PA} or proxy), there were three stocks which produced a False Alarm i.e. were considered by IUCN to be threatened with extinction, but the stock assessments indicate a biomass which is greater than the PA reference point. These were Golden redfish (*Sebastes norvegicus*) in subareas 5, 6, 12 and 14 (Iceland and Faroes grounds. West of Scotland. North of Azores. East of Greenland); Blue ling (*Molva dypterygia*) in subareas 6a, 7 and Division 5.b (Celtic Seas, English Channel and Faroes grounds); and Turbot (*Scophthalmus maximus*) in Subarea 4 (North Sea).

6) In terms of fishing mortality (F_{MSY} or proxy), there were 4 stocks which produced a False Alarm, i.e. were considered by IUCN to be threatened with extinction, but the stock assessments indicate a fishing mortality which was less than the MSY reference point. These were: Blue ling (Molva dypterygia) in subareas 6a, 7 and Division 5.b (Celtic Seas, English Channel, and Faroes grounds); Porbeagle (Lamna nasus) in the Northeast Atlantic; Porbeagle (Lamna nasus) in the Northwest Atlantic; and Turbot (*Scophthalmus maximus*) in Subarea 4 (North Sea).

7) In most cases these False Alarms can be attributed to the differences in timing of the assessments, with more recent and different information becoming available to the more recent fisheries assessments. In the one case

(porbeagle), the high uncertainty of an ICCAT assessment resulted in a False Alarm that may not be justified, and it could be argued that both systems recognise that the stock or species is in trouble (true positive).

8) In relation to commercially exploited fish, the two systems can be seen as complementary graduated indicators of status, with the fisheries assessments representing the first level of concern. The IUCN criteria could be appended as a matter of course to stock assessments where it is indicated that the stock is overfished or subject to overfishing.

Introduction

The status of many commercially exploited marine fish populations in the North East Atlantic has improved since the turn of the century due to improvements in management (Fernandes and Cook 2013). However, many species which are not commercially exploited, or those that are large (>150 cm), particularly in the taxonomic class of sharks and rays, are still under threat of extinction (Fernandes *et al*. 2017). Most of the commercially exploited stocks in the region are assessed more or less annually by the International Council for the Exploration of the Sea (ICES). Some of the large pelagic species (tunas, scombrids, and sharks) are assessed by the International Commission for the Conservation of Atlantic Tunas (ICCAT), although these assessments may be less frequent than those at ICES. The rest of the fish species, particularly those that are not commercially exploited, are included in assessments of extinction risk carried out by the International Union for the Conservation of Nature (IUCN). This report aims to compare the most recent assessments of these three systems for fish in the north east Atlantic.

ICES is a global organization which co-ordinates scientific activity and provides advice in support of the sustainable use of the oceans. As part of its work on fisheries, it collates data, and provides facilities for expert groups, to conduct assessments of stocks of fish and shellfish in the North East Atlantic (NEA). It then provides advice on how to conserve these stocks, typically by setting Total Allowable Catches (TACs) which align with the principle of Maximum Sustainable Yield[1] (MSY) within a Precautionary Approach[2] (PA) (ICES 2013). This advice forms a major part of the rules for managing European fishing fleets under the European Commission (ECs) Common Fisheries Policy (CFP).

ICES conducts its assessments on the basis of a stock which it defines as "*A part of a fish population usually with a particular migration pattern, specific spawning grounds, and subject to a distinct fishery. In theory, a Unit Stock comprises all the individuals of fish in an area, which are part of the same reproductive process. It is self-contained, with no emigration or immigration of individuals from or to the stock. On practical grounds, a fraction of the unit stock is considered a 'stock' for management purposes (or a management unit), as long as the results of the assessments and management remain close enough to what they would be on the unit stock.*" ICES stock assessments estimate four quantities: i) the total catch from the stock; ii) the number of young fish entering the stock (recruits); iii) the exploitation rate (Fishing mortality, F); and iv) the biomass of the adults in the stock (Spawning Stock Biomass, SSB). The status of a fish stock is then considered by comparing the most recent estimates of F and SSB against the respective reference points. ICES uses three types of reference points to assess stock status: i) Precautionary (PA, F_{PA} & B_{PA}); ii) Maximum Sustainable Yield (MSY, F_{MSY} & MSY $B_{trigger}$); and iii) Management Plan (MGT, F_{MGT} & B_{MGT}). A set of pictograms are used to represent the status of the stocks and their exploitation relative to each of these types of reference point where available. In the absence of reference points ICES presents a qualitative evaluation depending on data availability. There is no single concluding statement of the status of a stock: if data are available for a full stock assessment and all reference points are available, then

---

[1] The largest average catch or yield that can continuously be taken from a stock under existing environmental conditions.

[2] The ICES interpretation of precautionary approach is one whereby fisheries management keeps stock(s) adult biomass above the precautionary biomass reference point (B_{pa}) and the fishing mortality below the precautionary fishing mortality reference point (F_{pa}).

the status is described according to all six conditions (F in relation to F_{PA}, F_{MSY} and F_{MGT}; SSB in relation to B_{PA}, MSY $B_{trigger}$ and B_{MGT}).

ICCAT is an inter-governmental fishery organization responsible for the conservation of tunas and tuna-like species in the Atlantic Ocean and its adjacent seas. It provides a similar function to ICES in relation to fisheries, advising the EC on the sustainable management of tuna stocks. ICCAT conduct similar stock assessments to ICES, but report differently. Firstly, they use only the one set of reference points (MSY). Secondly, they summarise stock status in relation to F (stating whether the stock is subject to overfishing or not) and SSB (stating whether the stock is overfished or not), which is similar to the system in the USA [e.g. see Ianelli *et al*. (2014)]. Finally, they provide a pie chart with percentages assigned to each of three possible states, based on the uncertainty in the assessment. The three possible states are: i) overfished and overfishing; ii) overfished or overfishing; and iii) neither overfished nor overfishing (or "Kobe targets achieved").

The IUCN is "a membership Union uniquely composed of both government and civil society organisations. It provides public, private and non-governmental organisations with the knowledge and tools that enable human progress, economic development and nature conservation to take place together." One of IUCN's most well-known products is a conservation database called the IUCN Red List of Threated Species (henceforth, the Red List). This is an inventory of the global conservation status of plant and animal species, with an assessment of the risk of their extinction. To assess this risk the IUCN Red List categories and criteria are applied (IUCN 2012). These criteria are typically applied to species, although they can also applied to more general "taxon", and more specifically, to species at a regional level (IUCN 2012). There are nine IUCN Red List categories: Extinct (EX), Extinct in the Wild (EW), Critically Endangered (CR), Endangered (EN), Vulnerable (VU), Near Threatened (NT), Least Concern (LC), Data Deficient (DD) and Not Evaluated (NE). Two additional categories, Regionally Extinct (RE) and Not Applicable (NA), are used in regional Red List assessments. Species are classed as threatened if they fall within the categories CR, EN or VU. To classify as threatened, one or more of five quantitative criteria (A–E) related to population reduction (criterion A), geographic range (criterion B), population size and decline (criterion C), very small or restricted population (criterion D), and probability of extinction (criterion E) are examined for each species. Separate thresholds then allocate species to the individual categories based on the risk of extinction, with CR indicating an extremely high risk, EN a very high risk and VU a high risk. The NT category is for those species close to qualifying, or likely to qualify in future, as threatened. The LC category has a low risk of extinction.

The purpose of this report is to compare these three systems in respect of the status of fish stocks in the NEA. Specifically, it is to compare recent regional IUCN listings for key EU stocks, with corresponding information from stock assessments carried out by ICES and ICCAT. The list of ICES stocks were those provided in their stock assessment database[3] (and see methods Section 2.2); ICCAT stock assessments were found in ICCAT (2018).

[3] http://standardgraphs.ices.dk/stock-List.aspx

Specified task

"Compare the Red list categories from 2015 IUCN listing of species of European marine fish with the most recent, up to 2017, estimate of status (B_{PA} (or proxy) and F_{MSY} (or proxy)) of fish stocks assessed by ICES and ICCAT. "

Methods

IUCN data

Data were taken from the European Red List of Marine Fishes (Nieto *et al*. 2015, Fernandes *et al*. 2017). Assessments of extinction risk were carried out over a series of workshops from 2013 to 2014, for1022 species, although 202 were data deficient. 67 species were designated as threatened (see Appendix 1), and, of these, 21 were Critically Endangered (CR), 23 were Endangered (EN) and 23 were Vulnerable (VU). A further 26 species were considered Near Threatened (NT). The vast majority of species (71.1%, 725 species) were considered to be Least Concern (LC). A database of the results of the assessment was made available from IUCN. There were 6 fields in the database: taxonomic Class, Order, Family and Species; IUCN red list category and criteria.

ICES data.

The ICES stock assessment database contains data from the latest year of assessments. A full list of the 267 stocks for which ICES attempts assessments was obtained from ICES (Inigo Martinez pers. Comm.) Downloads were made of the database from 2018, 2017, 2016, 2015 and 2014, to ensure that assessments of all stocks were available, because each year's dataset only contains the assessments carried out in that year: some stocks are assessed every two years and others over longer time periods. In the current case, of the 267 stocks listed, data were available for 252 stocks: those missing were not available in the 2014, 2013, 2012 or 2011 datasets posted on the ICES website. Most reference points were available in the databases.

Where B_{PA} reference points were not available, the MSY or MGT reference point was used. Where F_{MSY} reference points were not available, the PA or MGT reference point was used. There were 3 examples were F reference points were not available, but were available as harvest rates on the ICES advice sheets: these were for the Icelandic stocks of haddock, saithe and cod. These were converted according to $F = -\ln(1-(HR_{ref\ point}))$ where HR = Harvest Rate. Some taxon were listed by their genus (Lepidorhombus, Ammodytes, *Beryx*): for the purposes of this comparison, these were converted to the principal species *Lepidorhombus whiffiagonis, Ammodytes marinus and Beryx splendens*. The turbot, classed by ICES as *Psetta maxima*, was converted to *Scophthalmus maximus* (as used by IUCN).

ICCAT data

ICCAT data for 16 of the specified stocks were taken from ICCAT (2018). Reference points were not always available for the ICCAT assessments but the ratios of B/B_{MSY} and F/F_{MSY} were, as well as the stock status. ICCAT report status in terms of an overall conclusion as to whether the stock is overfished or not, and whether it is subject to overfishing (see Section 1 above). They also report a range of values of F/F_{MSY} and B/B_{MSY} which often encompassed 1. In such cases the average value of the range was taken, being careful to ensure that they also coincided with the stock status definitions given in ICCAT (2018).

The most contentious set of values concerned the assessment of Porbeagle in the North East Atlantic. In terms of F/F_{MSY} the reported range was 0.04 to 3.45. The relevant footnote for this states "*Range obtained from BSP (low) and ASPM (high) models. The value of 0.04 from the BSP corresponds to a biologically unrealistic scenario; all results from the BSP scenarios ranged from 0.70 to 1.26.*" Given that the stock is not subject to "overfishing", so the ratio should be less than 1, the average of 0.7 and 1.26 was taken for F/F_{MSY} (i.e. 0.98). In terms of B/B_{MSY} the reported range was 0.09 to 1.93 and the footnote states "*Range obtained from BSP (high) and ASPM (low) models. Value from ASPM model is SSB/SSB*$_{MSY}$*. The value of 1.93 from the BSP corresponds to a biologically unrealistic scenario; all results from the other BSP scenarios ranged from 0.29 to 1.05.*" Given that the stock is "overfished", so the ratio should be less than 1, the average of 0.29 and 1.05 was taken for B/B_{MSY} (i.e. 0.67).

The combined ICES and ICCAT data set

The final dataset combined the available information from ICES and ICCAT and consisted of 268 stocks of 87 species. Status (B_{PA} or proxy) reference points were available for 98 stocks. F_{MSY} (or proxy) reference points were available for 108 stocks. These include those inferred reference points from the values of B/B_{MSY} and F/F_{MSY} for the ICCAT stocks. Assessments were carried out from 2018 to as far back as 2008 (e.g. Northeast Atlantic Porbeagle).

Fisheries assessments and conservation status comparisons.

The IUCN threat category (threatened or not) was compared with the two ratios F/F_{MSY} and B/B_{MSY} reflecting fishing mortality and status as defined by ICES or ICCAT in the specified task (Section 1.1). To be considered sustainable, the ratios F/F_{MSY} and B/B_{MSY} had to be less than, and greater than 1, respectively, for fishing mortality and status (as defined in the ICES request). All stocks were compared to the single conservation status attributed to the species. Four possible outcomes of this analysis were defined with respect to the IUCN definition of the threat of extinction (Rice 2003, Davies and Baum 2012):

1. True positive, where the two systems concur, and where a stock is not sustainable and the threat criteria are met;
2. True negative, where the two system concur, and where a stock is sustainable and the threat criteria are not met;
3. Misses, where a stock is exploited unsustainably but does not meet the threat criteria;
4. False alarms, where the stock is exploited sustainably but the threat criteria are met.

Results

Comparison with B_{PA}

In terms of status (B_{PA} or proxy), there were three stocks which produced a False Alarm i.e. were considered by IUCN to be threatened with extinction, but the stock assessments indicate a biomass which is greater than the PA reference point (Figure 1). These were:

i. Golden redfish (*Sebastes norvegicus*) in subareas 5, 6, 12 and 14 (Iceland and Faroes grounds. West of Scotland. North of Azores. East of Greenland), where:
    a. B_{PA}=220,000 t
    b. SSB = 342,100 t
    c. B/B_{PA} = 1.555
    d. IUCN category = VU
ii. Blue ling (*Molva dypterygia*) in subareas 6â€"7 and Division 5.b (Celtic Seas. English Channel. and Faroes grounds), where:
    a. B_{PA}= 75,000 t
    b. SSB = 101,501 t
    c. B/B_{PA} = 1.35
    d. IUCN category = VU
iii. Turbot (*Scophthalmus maximus*) in Subarea 4 (North Sea), where:
    a. B_{PA}= not defined (relative assessment)
    b. B/B_{PA} = 1.36
    c. IUCN category = VU

Three stocks produced a true positive hit (i.e. IUCN considered the species to be threatened with extinction and the stock assessment indicated a biomass less than the PA reference point. These were:

i. Beaked Redfish (*Sebastes mentella*) in Subareas V. XII. XIV and NAFO Subareas 1+2 (Deep Pelagic stock > 500 m deep)
    a. B_{PA} = 1.09
    b. SSB = 0.48
    c. B/B_{PA} = 0.44
    d. IUCN category = EN
ii. Golden Redfish (*Sebastes norvegicus*) in Subareas I and II
    a. B_{PA} = not defined (relative assessment)
    b. SSB = not defined (relative assessment)
    c. B/B_{PA} = 0.25
    d. IUCN category = VU
iii. Spurdog (*Squalus acanthias*) in the Northeast Atlantic
    a. B_{PA} = not defined (B_{MSY} = 964563 t)
    b. SSB = 50988 t
    c. B/B_{PA} = 0.05
    d. IUCN category = EN

There were 20 stocks classed as a miss (IUCN did not class them as threatened, but stock assessments indicate biomasses lower than B_{PA}); and 58 stocks as true negative hits (IUCN not threatened, stock assessments biomass higher than B_{PA}).

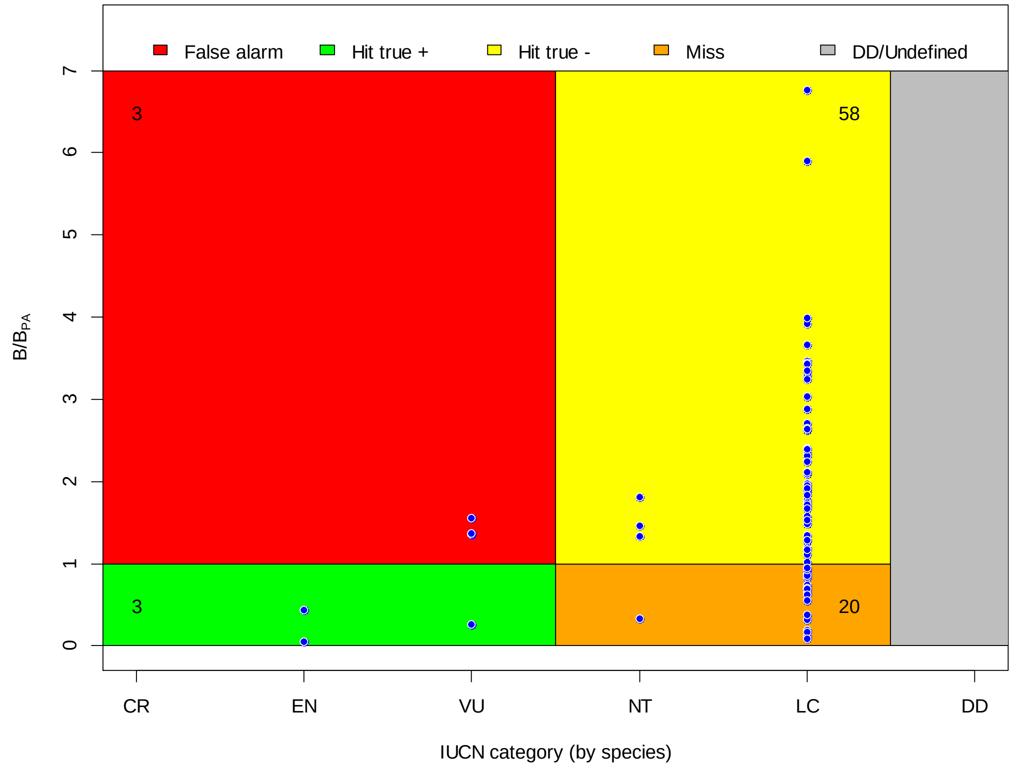


Figure 2. Performance of the IUCN Red List in relation to stock status as defined by the estimate of biomass relative to its precautionary reference point (B/B_{PA}, on y axis). Each point is a stock, with the species classified according to the threat criteria of the IUCN Red List (x axis). Red List categories are Critically Endangered (CR), Endangered (EN), Vulnerable (VU), Near Threatened (NT), Least Concern (LC) and Data Deficient (DD). Shading indicates: hits, where the two systems concur, either because the spawning stock biomass is below the reference point and the threat criteria are met (true positive, in green), or because a stock is above the reference point and the threat criteria are not met (true negative, in yellow); misses, in orange, where a stock is at low biomass but does not meet the threat criteria; and false alarms, in red, where the stock is at high biomass but the threat criteria are met. Numbers in each quadrant refer to the number of stocks. See text for further details.

Comparison with F_{MSY}

In terms of fishing mortality (F_{MSY} or proxy), there were 4 stocks which produced a false alarm (i.e. were considered by IUCN to be threatened with extinction, but the stock assessments indicate a fishing mortality which was less than the MSY reference point (Figure 2). These were:

i. Blue ling (*Molva dypterygia*) in subareas 6â€“7 and Division 5.b (Celtic Seas, English Channel, and Faroes grounds)
   a. F_{MSY}= 0.12
   b. F = 0.03
   c. F/F_{MSY} = .025
   d. IUCN category = VU

ii. Porbeagle (*Lamna nasus*) in the Northeast Atlantic
   a. F_{MSY} = not defined (relative assessment)

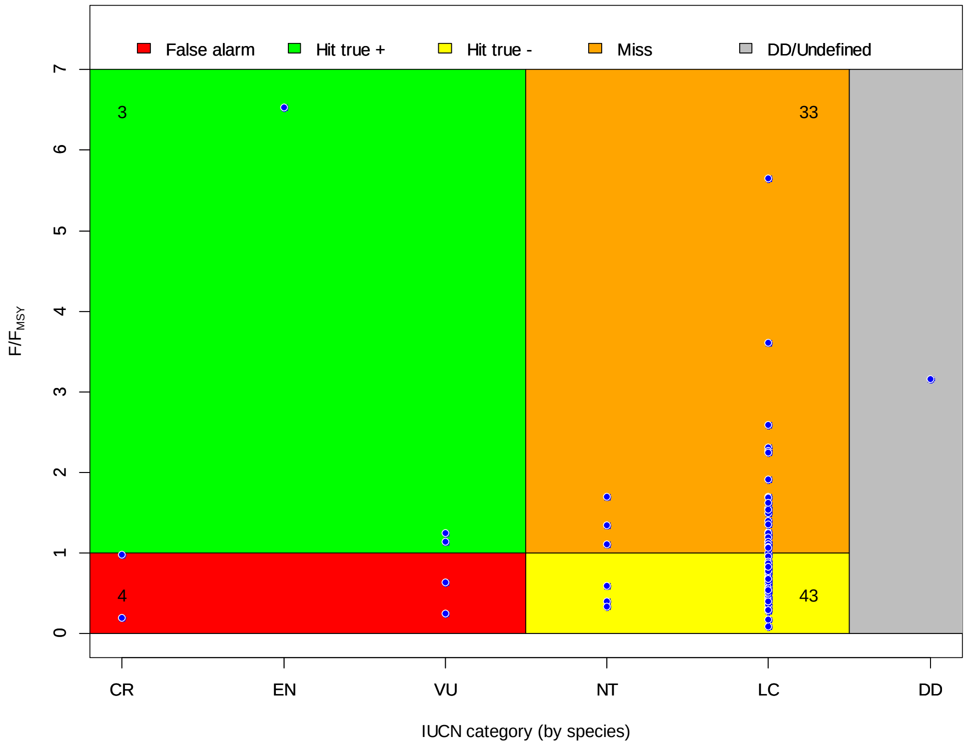


Figure 3. Performance of the IUCN Red List in relation to fishing mortality (F) as defined by the estimate of F relative to its MSY reference point (F/F_{MSY}, on y axis). Each point is a stock, with the species classified according to the threat criteria of the IUCN Red List (x axis). Red List categories are Critically Endangered (CR), Endangered (EN), Vulnerable (VU), Near Threatened (NT), Least Concern (LC) and Data Deficient (DD). Shading indicates: hits, where the two systems concur, either because the fishing mortality is above the reference point and the threat criteria are met (true positive, in green), or because the fishing mortality is below the reference point and the threat criteria are not met (true negative, in yellow); misses, in orange, where the fishing mortality is high but does not meet the threat criteria; and false alarms, in red, where the fishing mortality is low relative to the reference point, but the threat criteria are met. Numbers in each quadrant refer to the number of stocks. See text for further details.

b. F = not defined (relative assessment)
c. F/F_{MSY} = 0.98
d. IUCN category = CR

iii. Porbeagle (*Lamna nasus*) in the Northwest Atlantic
a. F_{MSY} = not defined (relative assessment)
b. F = not defined (relative assessment)
c. F/F_{MSY} = 0.195
d. IUCN category = CR

iv. Turbot (*Scophthalmus maximus*) in Subarea 4 (North Sea)
a. F_{MSY} = not defined (relative assessment)
b. F = not defined (relative assessment)
c. F/F_{MSY} = 0.63
d. IUCN category = VU

Three stocks produced a true positive hit (i.e. IUCN considered the species to be threatened with extinction and the stock assessment indicated a fishing mortality greater than the MSY reference point (Figure 2). These were:

These were:

i. Beaked Redfish (*Sebastes mentella*) in Subareas V. XII. XIV and NAFO Subareas 1+2 (Deep Pelagic stock > 500 m deep)
   a. F_{MSY} = 0.17
   b. F = 1.11
   c. F/F_{MSY} = 6.52
   d. IUCN category = EN
ii. Golden redfish (*Sebastes norvegicus*) in subareas 5, 6, 12 and 14 (Iceland and Faroes grounds. West of Scotland. North of Azores. East of Greenland).
   a. F_{MSY} = 0.097
   b. F = 0.11
   c. F/F_{MSY} = 1.14
   d. IUCN category = VU
iii. Golden Redfish (*Sebastes norvegicus*) in Subareas I and II
   e. F_{MSY} = not defined (relative assessment)
   f. F = not defined (relative assessment)
   a. F/F_{MSY} = 1.25
   b. IUCN category = VU

There were 33 stocks classed as a miss (IUCN did not class them as threatened, but stock assessments indicate fishing mortalities greater than F_{MSY}); and 43 stocks as true negative hits (IUCN not threatened, stock assessments fishing mortality lower than F_{MSY}).

Other observations

Yellowfin tuna (*Thunnus albacares*) were not assessed in the European Red List. Globally, they were classed as NT in 2011 (near threatened, which is not a threatened category). ICCAT consider them to be overfished, but not subject to overfishing (low biomass and low F relative to reference points). Using this assessment, the stock would be a miss in terms of status (B_{PA}) and a true negative hit in terms of fishing mortality (F_{MSY}).

Prawns (*Nephrops norvegicus*) were also not assessed by the European Red List [of marine fish]. Globally, they were classed as LC in 2009. There are numerous stocks and sub-stocks (functional units) of prawns in the ICES database, with several assessed with biomasses lower than the reference point.

Discussion

There are four significant distinctions between the assessments made by the IUCN and those made by the ICES and ICCAT fisheries stock assessments. The first and most important, as stated in Fernandes *et al*. (2017) is that "IUCN is concerned with extinction risk, whereas fisheries assessments are concerned with sustainable exploitation." The second is that IUCN work at the species as opposed to the stock level of fisheries assessments, even though, as in the current case, the European Red List exercise was regionalised to European marine waters. Nonetheless, there are situations where one species can encompass many stocks. The third distinction is that fisheries assessments, by their very nature, deal with a limited number of commercially exploited species. The ICES & ICCAT data included here had a total of 87 species (of 268 stocks), whereas the IUCN database consisted of more than ten times the number of species, and almost four times more species than the number of ICES & ICCAT stocks considered here (1022). Finally, and partly as a consequence of the latter point, fisheries assessments are typically conducted more frequently than the more numerous IUCN assessments. All of these factors can lead to discrepancies between the two systems, if indeed they should ever be considered equivalent. A stock being above a sustainable exploitation rate or below a precautionary reference biomass, does not necessarily equate to an extinction risk. Equally, if a fish stock is classified as sustainable [by whatever means], it may seem contradictory (though theoretically possible) for IUCN to place the species in a threatened category. These False Alarms are indicated in Figures 1 and 2 as red panels. It should be noted that they are "False Alarms" to the IUCN system, assuming the fisheries assessments are correct such that if the biomass is above the reference point there can be no extinction risk. This follows the framework used in (Davies and Baum 2012) which was based on the interpretation of signal detection theory defined in (Rice 2003). This assumes that there is a true signal being detected and distinguished from any noise in the measurement. Strictly speaking, there is no true signal in the use of this framework, because the two systems measure two different signals: IUCN measures extinction risk; ICES & ICCAT measure the risk of overfishing. The framework serves merely as a means to highlight differences between the two given that both extinction and overfishing can be considered signals worth detecting in order to employ some, mitigation/management action.

False alarms were detected in 3 stocks (Section 3.1): Golden redfish (Iceland – Azores); Blue ling (English channel to Faroes); and North Sea turbot. Golden redfish (*Sebastes norvegicus*) were classified by IUCN as VU under criteria A2bd (An observed, estimated, inferred or suspected population size reduction of ≥80% over the last 10 years or three generations). Further details for the species are not available on the Red List website. There are two stocks of Golden redfish in the ICES database. The stock that produced the false alarm was the one which occurs from Iceland to the Azores (coded as reg.27.561214). When the stock was assessed in 2017, there was an estimated SSB of 342,100 t which has been rising since the mid-1990s (Figure 3). Although ICES has changed the areas allocated to this stock in recent years, the previous stock (coded smr.5614) had similar trends in its assessment of 2014. It is puzzling to consider why this species was classified as VU. One possible explanation is the assessors may have only considered the stock in ICES areas I and II where the stock has been in relative decline since the mid-1990s (Figure 4).

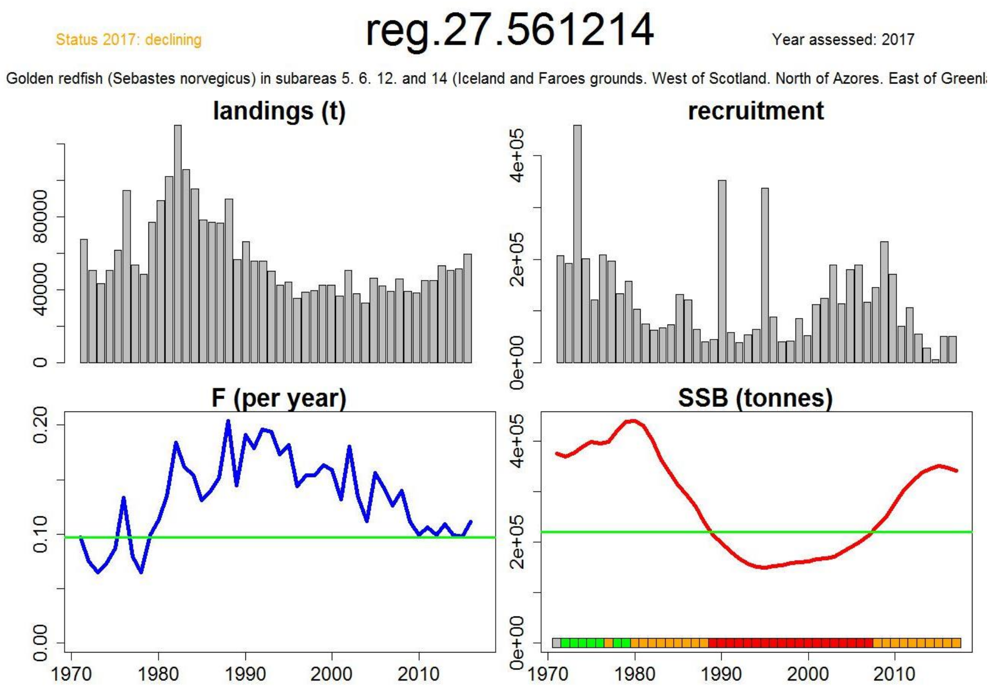


Figure 4. Stock assessment outputs for the stock of Golden redfish in ICES sub areas 5, 6, 12 and 14, showing landings in tonnes (top left), recruitment in thousands (top right panel), fishing mortality (bottom left) and Spawning Stock Biomass (SSB, in tonnes, bottom right). The colour coding on the bottom right panel refers to the status as defined by Fernandes *et al* (2017): sustainable (in green), overfished (red), recovering (yellow) and declining (orange).

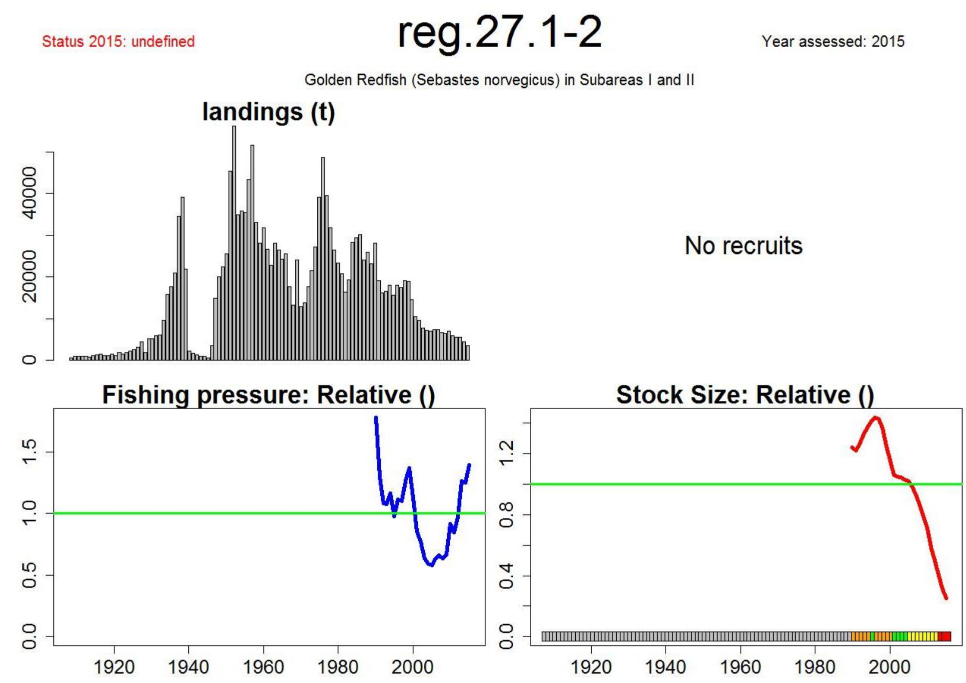


Figure 5. Stock assessment outputs for the stock of Golden redfish in ICES sub areas 1 and 2, showing landings in tonnes (top left), recruitment not estimates (in top right panel), fishing mortality (bottom left) and Spawning Stock Biomass (SSB, in tonnes, bottom right). The colour coding on the bottom right panel refers to the status as defined by Fernandes *et al* (2017): sustainable (in green), overfished (red), recovering (yellow) and declining (orange).

Blue ling is another stock which produced a false alarm. The fisheries stocks assessment determined that the stock in the Celtic Seas, Channel and Faroes of this species to be above the biomass reference point, but IUCN classed the species as VU. This is another stock which has a recent change in its designation (in 2017 this was included in a broader area), but more importantly, there were no biomass trends prior to 2018 for this stock. So when the IUCN assessment was carried out in 2013 and 2014 there was no information on biomass from ICES. North Sea turbot also produced a False Alarm. In this case the most recent ICES assessment is quite different from the ICES assessments of 2015, 2014 and 2013, which would have been available at the time of the IUCN assessment. The 2015 ICES assessment of North Sea turbot, for example, shows a long term decline in biomass, whereas the very next years' assessment [2016] shows an increase in biomass since the late 1990's (Figure 5). These False Alarms, or discrepancies, are, therefore, possible due to a mismatch in the timing of either assessment, or due to considerations of particular stocks rather than the species as a whole.

In terms of F_{MSY} four stocks produced a false alarm. Two of these have been discussed already in relation to B_{PA} (Blue ling and Turbot) and the same explanations apply (see above). The other two stocks were porbeagle (*Lamna nasus*) in the Northeast Atlantic and northwest Atlantic. IUCN classify porbeagle as CR. ICCAT's assessment of the northeast Atlantic is that F/F_{MSY} is between .04 and 3.45. They infer, however, that a more realistic range is 0.70 to 1.26, so a value of 0.98 was used here (see Section 2.3). This is an extremely uncertain estimate, and given that this is so close to 1, it is difficult to justify a significantly strong distinction between this and the IUCN assessment. On the basis of this uncertainty, a true positive hit was also possible, whereby both systems recognise that the stock/species is in trouble. The northwest Atlantic population is less contentious, as F/F_{MSY} was 0.195, but this is perhaps less relevant to Europe.

Given the large number of species in LC categories, it is more common for a discrepancy to occur whereby a species is not considered threatened but the stock may be overfished. A good example of this is the cod, *Gadus morhua*, classed as LC by IUCN. In Europe, the stock of Arctic cod alone is immense, with a spawning stock biomass of almost 2 million tonnes representing approximately 1.7 billion fish. Yet at the stock or sub-regional level, west of Scotland cod for example, has been in decline for decades and would almost certainly classify as Critically Endangered (CR) under Criterion A (reduction in population size based on an estimated population size reduction of ≥90% over the last 10 years or three generations, whichever is the longer). IUCN does not, however, regionalise to the level of the stock, in this, and indeed most cases. The larger the stock unit, the closer the two schemes are likely to be. So it is not surprising that the true positive hits, in biomass and fishing mortality comparisons, were stocks that encompass large regions (redfish and spurdog).

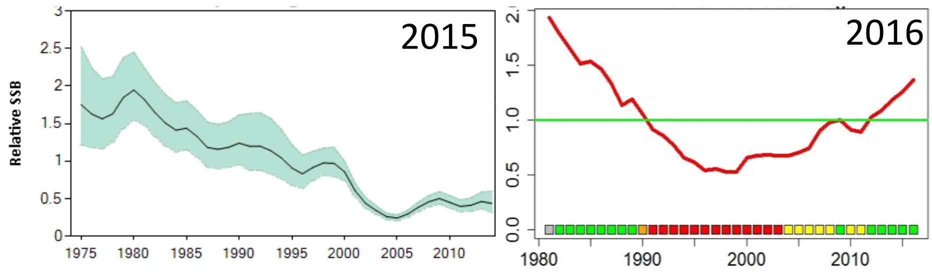


Figure 6. Relative spawning stock biomass times series for the stock of North Sea turbot (Scophthalmus maximus) as carried out by ICES in 2015 (left panel) and the most recent assessment in 2016 (right panel). The colour coding on the right panel refers to the status as defined by Fernandes *et al* (2017): sustainable (in green), overfished (red), recovering (yellow) and declining (orange).

There may be some merit in considering a hierarchy for employing the two systems, as it would be expected that overfishing would come before extinction. As Fernandes *et al* (2017) argue: "The two classification schemes can, therefore, be seen as complementary graduated indicators of status, with the stock sustainability [in their case by combining the two components of biomass and exploitation] representing the first level of concern. If a stock is overfished then further examination under the IUCN framework is merited to determine if there is an extinction risk." Given the different signals the two systems measure (overfishing by ICES & ICCAT and extinction risk by IUCN), the large number of "misses", as defined here (and in Davies and Baum, 2012), was expected. This is simply because an overfished stock does not necessarily equate to an extinction risk. There were many more "Misses" than "False Alarms", and as most of the "False Alarms" were explained by mismatches in timing of the assessments, the ICES & ICCAT systems could be seen as more pessimistic than the IUCN Red List, if these are considered equivalent in assessing the status of fish. The simple fact is that ICES/ICCAT and IUCN measure different signals, and it would be valuable to assess both signals if possible.

Fernandes *et al* (2017) also note that "Conversely, if a species is deemed to have a low risk of extinction (LC), it is not to say that certain local stocks may not be at risk. However, as stock assessments are updated every year and IUCN Red List assessments are much less frequent, discrepancies may yet occur. An important feature of the IUCN system is that it can be applied to species for which there is no analytical stock assessment. So it may be pertinent for Red List assessments to be appended to stock assessments, particularly in cases where those stocks are overfished or where data are deficient (for example, in terms of reference points or fishing mortality)." It is unlikely, given the resources available to IUCN and the number of species involved, that their assessments will be conducted on an annual basis. However, as ICES assessments are conducted more regularly, and annually for the most important stocks, it could be feasible to apply IUCN criteria as a matter of routine, append these to the assessments, and update the Red List. This would solve any potential conflicts in designation for many commercial species. However, this would only solve the problem of assessment timing and the complementary use of information: the issue of stocks vs species would still remain, but could be mitigated by stipulating the stock definition in line with the IUCN's flexibility to apply their criteria to 'taxon'.

Appendix 1

List of European marine fish species listed as regionally threatened according to the Red List conducted by the International Union for Conservation of Nature. Cat = IUCN Red List Category, where CR=Critically Endangered, EN=Endangered; VU=Vulnerable. Criteria follow those of IUCN (2012).

Class	Order	Species	Cat	Red List Criteria
Actinopterygii	Acipenseriformes	*Acipenser gueldenstaedtii*	CR	A2bcde
Actinopterygii	Acipenseriformes	*Acipenser naccarii*	CR	A2bcde; B2ab(i,ii,iii,iv,v)
Actinopterygii	Acipenseriformes	*Acipenser nudiventris*	CR	A2cd
Actinopterygii	Acipenseriformes	*Acipenser stellatus*	CR	A2cde
Actinopterygii	Acipenseriformes	*Acipenser sturio*	CR	A2cde; B2ab(ii,iii,v)
Actinopterygii	Acipenseriformes	*Huso huso*	CR	A2bcd
Chondrichthyes	Lamniformes	*Carcharodon carcharias*	CR	C2a(ii)
Chondrichthyes	Lamniformes	*Lamna nasus*	CR	A2bd
Chondrichthyes	Lamniformes	*Carcharias taurus*	CR	C2a(ii)
Chondrichthyes	Lamniformes	*Odontaspis ferox*	CR	A2bcd
Chondrichthyes	Rajiformes	*Gymnura altavela*	CR	A2bd
Chondrichthyes	Rajiformes	*Pteromylaeus bovinus*	CR	A2c
Chondrichthyes	Rajiformes	*Pristis pectinata*	CR	A2b; D
Chondrichthyes	Rajiformes	*Pristis pristis*	CR	A2b; D
Chondrichthyes	Rajiformes	*Dipturus batis*	CR	A2bcd+4bcd
Chondrichthyes	Rajiformes	*Leucoraja melitensis*	CR	A2bcd+3bcd
Chondrichthyes	Rajiformes	*Rostroraja alba*	CR	A2bd
Chondrichthyes	Squaliformes	*Centrophorus granulosus*	CR	A4b
Chondrichthyes	Squatiniformes	*Squatina aculeata*	CR	A2bcd
Chondrichthyes	Squatiniformes	*Squatina oculata*	CR	A2bcd+3cd
Chondrichthyes	Squatiniformes	*Squatina squatina*	CR	A2bcd+3d
Actinopterygii	Cyprinodontiformes	*Aphanius iberus*	EN	A2ce
Actinopterygii	Gadiformes	*Coryphaenoides rupestris*	EN	A1bd
Actinopterygii	Perciformes	*Anarhichas denticulatus*	EN	A2b
Actinopterygii	Perciformes	*Epinephelus marginatus*	EN	A2d

Class	Order	Species	Cat	Red List Criteria
Actinopterygii	Perciformes	*Pomatoschistus tortonesei*	EN	B2ab(ii,iii)
Actinopterygii	Scorpaeniformes	*Sebastes mentella*	EN	A2bd
Chondrichthyes	Carcharhiniformes	*Carcharhinus longimanus*	EN	A2b
Chondrichthyes	Carcharhiniformes	*Carcharhinus plumbeus*	EN	A4d
Chondrichthyes	Lamniformes	*Alopias superciliosus*	EN	A2bd
Chondrichthyes	Lamniformes	*Alopias vulpinus*	EN	A2bd
Chondrichthyes	Lamniformes	*Cetorhinus maximus*	EN	A2abd
Chondrichthyes	Rajiformes	*Mobula mobular*	EN	A2d
Chondrichthyes	Rajiformes	*Leucoraja circularis*	EN	A2bcd
Chondrichthyes	Rajiformes	*Raja radula*	EN	A4b
Chondrichthyes	Rajiformes	*Glaucostegus cemiculus*	EN	A3bd
Chondrichthyes	Rajiformes	*Rhinobatos rhinobatos*	EN	A2b
Chondrichthyes	Squaliformes	*Centrophorus lusitanicus*	EN	A4b
Chondrichthyes	Squaliformes	*Centrophorus squamosus*	EN	A4b
Chondrichthyes	Squaliformes	*Deania calcea*	EN	A4d
Chondrichthyes	Squaliformes	*Dalatias licha*	EN	A3d+4d
Chondrichthyes	Squaliformes	*Echinorhinus brucus*	EN	A2bcd
Chondrichthyes	Squaliformes	*Centroscymnus coelolepis*	EN	A2bd
Chondrichthyes	Squaliformes	*Squalus acanthias*	EN	A2bd
Actinopterygii	Beryciformes	*Hoplostethus atlanticus*	VU	A1bd
Actinopterygii	Clupeiformes	*Alosa immaculata*	VU	B2ab(v)
Actinopterygii	Gadiformes	*Molva dypterygia*	VU	A1bd
Actinopterygii	Perciformes	*Mycteroperca fusca*	VU	B2ab(v)
Actinopterygii	Perciformes	*Bodianus scrofa*	VU	B2ab(iv,v)
Actinopterygii	Perciformes	*Labrus viridis*	VU	A4ad
Actinopterygii	Perciformes	*Umbrina cirrosa*	VU	A2bc
Actinopterygii	Perciformes	*Orcynopsis unicolor*	VU	A2bde
Actinopterygii	Perciformes	*Dentex dentex*	VU	A2bd
Actinopterygii	Pleuronectiformes	*Hippoglossus hippoglossus*	VU	A2ce
Actinopterygii	Pleuronectiformes	*Scophthalmus maximus*	VU	A2bd

Class	Order	Species	Cat	Red List Criteria
Actinopterygii	Salmoniformes	*Salmo salar*	VU	A2ace
Actinopterygii	Scorpaeniformes	*Sebastes norvegicus*	VU	A2bd
Chondrichthyes	Carcharhiniformes	*Galeorhinus galeus*	VU	A2bd
Chondrichthyes	Carcharhiniformes	*Mustelus mustelus*	VU	A2bd
Chondrichthyes	Carcharhiniformes	*Mustelus punctulatus*	VU	A4d
Chondrichthyes	Rajiformes	*Dasyatis centroura*	VU	A2d
Chondrichthyes	Rajiformes	*Dasyatis pastinaca*	VU	A2d
Chondrichthyes	Rajiformes	*Myliobatis aquila*	VU	A2b
Chondrichthyes	Rajiformes	*Leucoraja fullonica*	VU	A2bd
Chondrichthyes	Rajiformes	*Raja maderensis*	VU	D2
Chondrichthyes	Squaliformes	*Centrophorus uyato*	VU	A2b
Chondrichthyes	Squaliformes	*Oxynotus centrina*	VU	A2bd

References

Davies, T. D. and J. K. Baum (2012). "Extinction risk and overfishing: reconciling conservation and fisheries perspectives on the status of marine fishes." *Scientific Reports* **2**.

Fernandes, P. G. and R. M. Cook (2013). "Reversal of fish stock decline in the northeast Atlantic." *Current Biology* **23**(15): 1432-1437.

Fernandes, P. G., G. M. Ralph, A. Nieto, M. García Criado, P. Vasilakopoulos, C. D. Maravelias, R. M. Cook, *et al*. (2017). "Coherent assessments of Europe's marine fishes show regional divergence and megafauna loss." *Nature Ecology and Evolution* **1**: 0170.

Ianelli, J., T. Honkalehto, S. Barbeaux and S. Kotwicki (2014). "Assessment of the walleye pollock stock in the Eastern Bering Sea." *In: Stock assessment and fishery evaluation report for the groundfish resources of the Bering Sea/Aleutian Islands regions. Alaska Fisheries Science Center, National Marine Fisheries Service, Seattle, WA,*: 55-156.

ICCAT (2018). "ICCAT report for the biennial period, 2016-17." *PART I (2017)* **Vol. 2**: 427.

ICES (2013). "Report of the ICES Advisory Committee 2013." *ICES Advice, 2013* **Books 1-11**: 2095.

IUCN (2012). *Guidelines for application of IUCN Red List Criteria at regional and national levels: Version 4.0*. Gland, Switzerland and Cambridge, UK, IUCN.

IUCN (2012). *IUCN Red List Categories and Criteria: Version 3.1. Second edition.* Gland, Switzerland and Cambridge, UK, IUCN Species Survival Commission.

Nieto, A., G. M. Ralph, M. T. Comeros-Raynal, J. Kemp, M. G. Criado, D. J. Allen, N. K. Dulvy, *et al*. (2015). *European red list of marine fishes*, Publications Office of the European Union.

Rice, J. (2003). "Environmental health indicators." *Ocean & Coastal Management* **46**(3): 235-259.

Annex 2 Comparison of IUCN categories and fishery management reference points

September 2018

Gavin Fay[1]

[1]Department of Fisheries Oceanography, School for Marine Science and Technology, University of Massachusetts Dartmouth, 836 S Rodney French Boulevard, New Bedford, MA 02744, USA. gfay@umassd.edu, Tel: +1-508-910-6363

1. Summary

Using a Management Strategy Evaluation framework, comparisons were made between IUCN categories of conservation status and ICES fisheries reference points. A series of stock projection simulations were conducted for four species/stocks, western Atlantic Bluefin tuna (*Thunnus thynnus*), blue shark (*Prionace glauca*), Atlantic cod (*Gadus morhua*) in the Baltic Sea, and Red (blackspot) seabream (*Pagellus bogaraveo*).

Simulations were conditioned on available information for the four stocks, and a simulated pseudo-assessment was used in conjunction with an ICES control rule to project the stocks for 100 years under these management, subject to process error and estimation uncertainty. Simulations were compared by using simulated stock assessment estimates of spawning biomass relative to ICES biomass reference points, these same estimates for the magnitudes of population decline relative to thresholds associated with IUCN categories under criterion A. The distribution of stock outcomes from 1000 simulations for each species was also used to compute probabilities of extinction for assessment under IUCN criterion E.

Results suggested the ICES control rule was sufficient to maintain stocks above collapse thresholds, even in data-poor cases. IUCN listings occurred frequently when assessing historical trends in biomass for these stocks, all of which are currently thought to be at low levels (to varying extents). Listings of Vulnerable tended to become more frequent towards the end of stock projections when populations were on average at higher levels and moving around biomass targets. There was a lack of correspondence between IUCN category listings and estimates of stock status relative to reference points when stocks were estimated to be below reference points, particularly during stock recovery. Populations that began management projection periods at higher stock levels showed greater correspondence between IUCN category listings and stock status below to reference points, as revealed through inspection of confusion matrices (e.g. blue shark). When stock status was estimated to be above fisheries reference points, IUCN assessments also tended to give a positive result (i.e. decisions not to list stocks).

Analyses with Baltic cod demonstrate a problem with summing over multiple populations to compute estimates of the magnitude of population change – listing decisions under IUCN criteria A were insensitive to the status of 1 of 2 stocks relative to biomass reference points when the IUCN listing thresholds were applied at the aggregated multi-stock level.

The results suggest that MSE is a useful approach for evaluating the performance of metrics for determining population status, and could be used to generate proababilistic estimates of population collapse. More complicated operating models that include

changes in population dynamics at small stock size would provide more realistic estimates. Additional dimensions of uncertainty could be explored for the stocks considered to further disentangle the factors driving variability in outcomes in the analyses.

2. Methods

2.1. General Approach

Using a Management Strategy Evaluation (MSE, e.g. Bunnefeld *et al*. 2009) framework, a series of stock projection simulations were conducted for four species/stocks, western Atlantic Bluefin tuna (Thunnus thynnus), blue shark (Prionace glauca), Atlantic cod (Gadus morhua) in the Baltic Sea, and Red (blackspot) seabream (Pagellus bogaraveo). In each case, an operating model was developed from current understanding of stock characteristics and conditioned on available information and catch history to represent a simulated 'truth' of population dynamics. Operating models were projected forward in time for 100 years, with catches informed by a pseudo-stock assessment and a harvest control rule intended to mimic the behaviour of ICES control rules. Stock projections were therefore subject to both process error in the population dynamics (recruitment variability) and estimation error (of stock status and reference points). 1,000 simulations were conducted for each species and operating model scenario. In each year of the assessment projection, the estimated current spawning stock biomass was compared to ICES reference points, and the magnitude of population decline over the previous 3 generations was calculated (used by IUCN to assess categories under criterion A). The frequency of years and simulations during which stocks would have been assessed to meet listing criteria, as well as status relative to ICES biomass reference points (e.g. Bpa) were then summarized for each species/stock and scenario. In addition, extinction probabilities (for certain thresholds of relative population size) were calculated from the simulations by considering the ensemble of 1,000 simulations as stochastic projections given populations/stocks subject to management (thus enabling comparison of population simulations under IUCN criterion E).

2.2. Operating models

The operating models consisted of an age-structured, sex-specific population dynamics model, exploited by a single fishing fleet (i.e. including commercial and recreational landings, as well as discards), with fishing assumed to take place in the middle of the year. Full technical specifications for the generalised version of the operating model are detailed in Fay *et al*. (2011). This operating model has been used extensively to evaluate the performance of assessment methods and management strategies (e.g. Fay *et al*. 2011; Little *et al*. 2014; Klaer *et al*. 2012; Fay and Tuck 2011). Life history and stock-recruitment parameter values, and the historical time series of catches were taken from recent stock assessment reports and ICES catch databases (Table 1, Figs 1-4). Populations were assumed to be unexploited with equilibrium age structure in the initial year of the simulations – stock status in the terminal year of the historical period (i.e. at the beginning of the application of the management strategies) was determined by pre-specifying the relative spawning stock depletion (spawning biomass relative to unfished spawning biomass), and solving for historical population size given the known time series of historical catch and a random (for each simulation) vector of recruitment residuals. Additional operating model details specific to each stock are outlined below, and are summarized in Table 1 and Figs 1-4.

The operating models were projected forward in time for 100 years, with each year of the projection period comprising a pseudo-stock assessment, an application of the results of the stock assessment to a harvest control rule to obtain a total allowable catch (TAC) for that year, and then the implementation of that TAC in the operating model to update the population dynamics.

2.3. Pseudo-stock assessment

To approximate stock assessments for the four species, an estimate of current spawning biomass was generated from the operating model, given a specified level of uncertainty meant to reflect the accuracy of assessment methods for the species. As biased in stock assessments tend to persist over time (because consecutive assessments often use very similar data sets and methods), the assessment estimates were assumed to be correlated through time. Pseudo-assessment estimates of spawning biomass were generated following methods similar to Punt *et al*. (2008) and Wiedenmann *et al*. (2015):

$$\widehat{SSB_y} = SSB_y e^{\delta_y - \sigma^2/2};\ \ \delta_y = \rho\delta_{y-1} + \sqrt{1-\rho^2}\varphi_y;\ \ \varphi_y \sim N(0,\sigma^2)$$

where $\widehat{SSB_y}$ is the assessment estimate of spawning biomass in year y of the simulation, SSB_y is the operating model (true) spawning biomass, ρ determines the degree of autocorrelation in the assessment estimates between years, and σ^2 determines the imprecision of the assessment estimates. For the analyses here, ρ was set at 0.707 for all species (cf. Punt *et al*. 2008) and σ was set at 0.2 for the data-rich stocks (Bluefin tuna and Baltic cod) and 0.4 for the data-poor stocks (blue shark and blackspot bream).

Note that this method assumes a stock assessment estimate of abundance was available for all stocks, though in practice this estimate of stock status may be derived from different assessment methodologies.

2.4. Control Rule and estimated biological reference points

The TAC in each year of the projection period was calculated by comparing the pseudo-stock assessment estimate of spawning biomass to an estimate of MSY $B_{trigger}$ to calculate the fishing mortality rate to be applied, as per the methods used by ICES (2017). The control rule used was:

$$F_y = \begin{array}{ll} F_{MSY} & if\ \widehat{SSB_y} \geq MSY\ B_{trigger} \\ F_{MSY}\widehat{SSB_y}/MSY\ B_{trigger} & if\ \widehat{SSB_y} < MSY\ B_{trigger} \end{array}$$

This control rule assumes a linear ramp on the implemented fishing mortality from the estimate of F_{MSY} when the stock is estimated to be below MSY $B_{trigger}$.

The above control rule requires estimates of F_{MSY} and MSY $B_{trigger}$. For the simulations, these were estimated from the operating model in the initial year of the projection period, by assuming that MSY $B_{trigger}$ = 0.5B_{MSY}, with deterministic estimates of B_{MSY} and F_{MSY} being obtained given simple yield-per-recruit analysis with assumed M constant over age (in the OM M varied by age for several of the stocks, Table 1), and expected recruits calculated given the known (true OM value) for steepness of the stock-recruit relationship and average unfished recruitment. The estimates for these reference points used in the simulations are thus close to but are not exactly the same as the true OM values. The values of these reference points that are currently used in the most recent stock assessments for the four stocks were therefore not used in the simulations, and were estimated within the simulations such that the values for reference points are internally consistent with the simulated stock dynamics. The approach used to define F_{MSY} and MSY $B_{trigger}$ reference points is not identical to the approach used to define these quantities by ICES (which is stochastic), but is computationally feasible and captures the intent of these reference points (indeed, the assumption for MSY $B_{trigger}$ matches the ICES approach for stocks estimated using biomass dynamics models). $0.5B_{MSY}$ is also a commonly used fisheries reference point for defining stocks that are overfished when below this threshold.

After the value for F_y is calculated, this is then used with the current estimate of spawning biomass to calculate the total catch over ages that would be expected given this value of fishing mortality to define the TAC:

$$TAC_y = \frac{\widehat{SSB_y}}{SSB_y} \sum_{l=1}^{N_L} \sum_{s=1}^{2} w_{s,l} \sum_{a=0}^{A} \Phi_{l,s,a} N_{s,a,y}\, e^{-0.5M_a}(1 - e^{-S_a F_y})$$

where S_a is the selectivity at age a, $w_{s,l}$ is the weight at length bin l for sex s, $\Phi_{l,s,a}$ is the proportion of fish of sex s and age a in length bin l, N and M are the numbers and mortality at age, and N_L is the number of length bins, and A is the maximum age. This approach therefore assumes that the catch is calculated given the correct (OM) age structure, but does account for the effect of assessment error on the TAC given a value of F_y from the control rule.

2.5. Performance metrics

For each simulated assessment, two sets of performance metrics are calculated: (a) whether or not at that time point a stock would be estimated to meet IUCN category listing under criterion A, and (b) whether or not the assessment estimate is below ICES biomass reference points (B_{pa} and B_{lim}).

For listings under IUCN criterion A, an estimate of the magnitude of population size reduction is needed, measured over the longer of 10 years or 3 generations. Generation time was computed for the four stocks using standard definition and the values for natural mortality and fecundity at age. Estimates of the magnitude of biomass change for each assessment year were computed by comparing the assessment estimate of spawning biomass in the current year to that at time $y\text{-}T_{gen}\text{+}1$. Listings were recorded as occurring if the magnitude of population size reduction exceeded thresholds of: ≥30% (Vulnerable under A2-4), ≥ 50% (Vulnerable under A1, Endangered under A2-4), ≥ 70% (Endangered under A1), ≥ 80% (Critically Endangered under A2-4), and ≥ 90% (Critically Endangered under A1). These thresholds were evaluated using both the estimated status of the stock and the actual OM values for spawning biomass.

Comparison to ICES biomass reference points was achieved by comparing the assessment estimates of SSB to the estimated reference points. Estimates for B_{pa} and B_{lim} were calculated from the estimate for MSY $B_{trigger}$ by assuming $B_{lim} = 0.3\ B_{MSY}$, and $B_{pa} = 1.4 \times B_{lim}$ (cf ICES 2017).

A third set of metrics, intended to evaluate the probability of extinction from a quantitative analysis (IUCN criteria 5), were calculated by considering the distribution of outcomes over simulations for each stock scenario. The frequency (of simulations) in which spawning biomass in the terminal year of the projection fell below 10%, 5%, and 1% of unfished spawning biomass was calculated. While these thresholds are not extinction per se, they do represent very low stock size.

3. Results

3.1. Western Atlantic Bluefin tuna

The operating model for western stock of Atlantic Bluefin tuna was conditioned using information from a recent implementation of Stock Synthesis for this stock (ICCAT 2017). The simulation scenarios considered an older age at maturity ogive (age of maturity ~9). At the start of the management period (based on 2016), the assumed depletion was 15% of unfished spawning biomass (Fig 5A). Simulations considered a single stock exploited solely by catches obtained from that stock area – i.e. no mixing of stocks or fishery catches with an eastern population were considered.

Operating model simulations of stock trajectory modelled a depletion from 1950, with some evidence of increasing population at the end of the historical time series (Fig. 5A). During the assessment and management projection period, stock simulations under the control rule showed a modest recovery of biomass on average.

Due to the initial depletion of biomass from 1950, operating model values for the magnitude of biomass change exceeded thresholds for IUCN listing under criterion A in a large number of cases during the historical period (Fig 5B). However, following implementation of the control rule, the number of assessments that triggered listing was greatly reduced, with almost no instances of a population decline of more than 50% during the projection period (Fig. 5B). In the final 50 years of the simulations, a number of simulations (~15%) experienced biomass declines of >30%, which would correspond to listing of vulnerable under criteria A2-A4. These declines arise as the stock begins to oscillate around the equilibrium target biomass associated with the control rule. This general pattern was also observed in the assessment estimates of population decline (which would actually be used as assessors would not have knowledge of true dynamics, here represented by the operating model), with a slightly higher frequency of listings (~20%) of Vulnerable under A2-A4 toward the end of the time period (Fig 5C).

In contrast to the low frequency of IUCN listings, western Bluefin tuna stocks were assessed to be below both B_{pa} and B_{lim} in almost all simulations at the beginning of the management period (Fig 5D). This is expected as the model was pre-specified to begin the management period at very depleted levels, reflecting current estimates of stock status. The stock remained below MSY $B_{trigger}$ and B_{pa} in the majority of simulations throughout the projection period (Fig 5D), reflecting the fact that the deterministic reference points used when evaluating stock status are mis-specified. The frequency at which the stock was estimated to be below the limit reference point B_{lim} declined as the management period proceeded, equilibriating at around 35% of the time by the end of the simulations. These results suggest that the ICES reference points more accurately reflect the depleted nature of the stock compared to the IUCN listings, as an assessment conducted mid-way through the projection period would rarely result in a listing under IUCN criterion A, whereas the fisheries reference points would suggest (and demonstrated from the operating model trajectory) that the stock is still below target levels.

This lack of correspondence between IUCN listings and identification of being below fisheries reference points can also be viewed via a confusion matrix, in which simulated assessments are compared against both methods of determining status (Fig 6) for different combinations of thresholds and reference points. The relative frequency of outcomes for whether a biomass decline was estimated to have exceeded a threshold resulting in IUCN listing under criteria A (e.g. >= 50%) was largely insensitive to whether that assessment had resulted in the stock being estimated to be above or below biomass reference points (e.g. B_{pa}). If there was complete correspondence between IUCN criteria A listings and estimates of being below reference points, then the matrices in Fig 6 would show display as red in both the lower left AND upper right quadrants. For Bluefin tuna, there is indication that estimates of biomass decline resulting in listings of Vulnerable (A1) or Endangered (A2-A4) would have always been identified by comparison of assessment estimates to biological reference points – estimated declines of >=50% only occurred when stocks were assessed to be below B_{lim} or below B_{pa} (Fig 6A, 6B).

There was virtually zero (1 out of 1000 simulations) probability where the OM spawning biomass fell below 10% of unfished spawning biomass during the simulations (Fig 7). The control rule therefore was able to sustain the population above these very low levels. The value for B_{lim} used in the simulations was 21% unfished biomass (although the magnitude of this changed by simulation, the specifications are such that this reference point remains the same across simulations when expressed relative to unfished biomass).

3.2 Blue shark

The OM for Atlantic blue shark was conditioned using information from the 2015 stock assessment document (ICCAT 2015). This included both historical catch estimates, biological information, as well as model settings akin to those used by ICCAT (2015) authors in a Stock Synthesis model. Historical unfished biomass was defined in 1968, although catches were very high in this year (Fig 2). Consequently initial depletion of the population during the historical period of the simulations is likely much higher than that actually observed over this time period. Based on these assessment results, the blue shark population was estimated to be around BMSY in the most recent year (2014). The simulations were conditioned assuming the stock was at 25% unfished biomass at the start of the projection period. An additional scenario was also run where the OM setting was changed to begin the population at 40% unfished biomass when the assessment and management cycle is first applied in the simulations.

The Operating model spawning biomass trajectory for blue shark shows a depletion in the historical time period (Fig 8A), with the stock remaining on average around the depletion level of 25% unfished spawning biomass (close to BMSY) during the projection period. Worm plots of individual simulated stock trajectories show that stocks actually fluctuate quite far from this central tendency of the simulation distribution during the management period (Fig 8A). This reflects the more uncertain assessment for this stock than for Bluefin tuna (assessment σof 0.4 compared to 0.2 for bluefin). Coincident with the initial drop in biomass during the historical period, OM rates of population change triggered listings thresholds of >=30% and >=50% during most simulations in this period (Fig 8B), and estimated rates of population decline resulted in listings of Vulnerable (A1-A4) and/or Endangered (A2-A4) in an appreciable number of simulations (25% of simulations for Vulnerable under A2-4 and 10% of simulations for Vulnerable under A1 or Endangered under A2-4) throughout the management period (Fig 8C), with the exception of a dip in the 2030s which was likely the consequence of the quick increase in population size at the beginning of the projection period (Fig 8A).

The blue shark stock was estimated to be below B_{pa} approximately 20% of the time during the projection period (Fig 8D), with very low frequencies of assessments estimating that the stock was below B_{lim}. Confusion matrices of the outcomes for blue shark (Fig. 9) show improved correspondence between IUCN listings and estimates of being below reference points compared to Bluefin tuna, with IUCN population decline thresholds of >=50% being exceeded in 36% of simulated assessments estimated to be below B_{pa} (Fig 9A) and 65% of simulated assessments estimated to be below B_{lim} (Fig 9B). As for bluefin tuna, there were very few instances where a >=50% decline threshold was exceeded when the stock was not assessed to be below biomass reference points.

Despite the increased uncertainty in assessment and wider range of stock outcomes, there were no instances of simulated stocks falling below 10% of unfished spawning biomass (Fig 10).

3.3. Baltic cod

Two scenarios were run for Baltic cod. Simulations were conducted for separate eastern and western Baltic cod stocks, assuming complete independence between stock dynamics and fisheries (i.e. no mixed catches, or rather, assuming that catches are correctly identified to population of origin). OMs were conditioned on recent assessments (ICES 2018, ICES 2015). Both eastern and western stocks were assumed to be at 12.5% of unfished spawning biomass at the beginning of the management period. Biological parameters for both eastern and western stocks were assumed to be the same. To evaluate the effect of conducting assessments against IUCN listing criteria at a multi-population level, the assessment estimates of biomass from the separate eastern and western analyses were summed over both stocks to obtain a time series of biomass estimates for the total for both stocks combined. The estimates of biomass change from these combined indices were then assessed against IUCN listing thresholds under criterion A in the same way as for the other stocks, that is SSB in the past was summed over both stocks and compared to the sum over both stocks in the present.

As the same biological parameters and fishery selectivity were used for both the eastern and western stocks of Baltic cod, the results and performance metric behaviour for the individual east and west simulations were identical. Stocks responded quickly to implementation of control rules during the management period, with increases in stock size to oscillate around a biomass associated with the estimated deterministic target (Figs 11, 14). High OM steepness (0.8) and recruitment variability (sigma R = 0.6) leads to wide intervals for the summary of stock trajectories, although the control rule is able to sustain the stocks above 10% unfished spawning biomass (as evidenced from the time series of extinction probabilities – Figs 13, 16).

OM historical time series of spawning biomass show population declines sufficient to trigger thresholds associated with the Highly Endangered category, although theoretical assessments conducted in the later parts of the historical time series following reduction of catches show a much lower frequency of listings of all categories under criterion A (e.g. Fig 11B). Following stock recovery during the management period, listings of Vulnerable under criterion A1 are made in about 20% of simulations based on both OM biomass trajectories and assessment estimates of spawning biomass decline (e.g. Fig 11B-C). Following an initial recovery from very low stock size (12.5% B0 is less than B_{lim} which was 18.3% in this case) Baltic cod stocks are estimated to be below B_{pa} in 35% of simulations in any given year, and below MSY Btrigger in ~50% of simulations (e.g. Fig 11D).

Confusion matrices of assessment outcomes for both IUCN listings and designations of being below biomass reference points for Baltic cod are intermediate between results for Atlantic Bluefin tuna and blue shark, with 12% of simulated assessments where stocks are thought to be below B_{pa} also having a >=50% biomass decline (e.g. Fig 12A). This percentage increases to 22% for the same population decline threshold when compared to assessments with estimates of the stock being below B_{lim} (e.g. Fig 12B).

When IUCN biomass decline statistics are calculated from the summed spawning biomass of both eastern and western assessment estimates, the number of simulations in which listing categories are triggered increases slightly (Fig 17), but this increase is marginal. However, the effect of assessing against IUCN categories under criteria A at this aggregated scale is clear when viewing the associated confusion matrices (Fig 18). Instances of exceeding the >= 50% threshold are only triggered when both the eastern and western stocks are assessed to be below the biomass reference points (both B_{pa} and B_{lim}, Fig 18A & 18B). If only one stock is assessed to be below the biomass reference point then the number of IUCN listings remains very low (or zero). Here, the larger eastern stock effectively masks biomass changes in the smaller western stock.

3.4. Blackspot seabream

The simulations for red (blackspot) were conditioned on the dynamics and fishery of the species in the Bay of Biscay, as detailed by Lorance (2011). Blackspot seabream were assumed to be heavily depleted at the beginning of the management period (10% of unfished spawning biomass), with the stock having been at low levels for about 20 years. Pseudo-assessments for blackspot seabream in the simulations were assumed to have a CV of 40%, reflecting a data poor status. When managed under the control rule, simulations showed stock recovery by 2040, with a wide 95% interval of spawning stock outcomes, as individual stock trajectories fluctuate around the biomass associated with the estimated reference points (Fig. 19A).

Mirroring the historical depletion of the blackspot seabream resource, OM biomass declines exceeded the Endangered and Critically Endangered thresholds right up until a few years before the management period (Fig. 19B). However, the number of assessment estimates of biomass decline once the management period began dropped to zero as the spawning stock recovered – even with the large CV on the pseudo-assessment estimates of spawning biomass the increase in stock size during this period was evidently plainly apparent. Towards the latter half of the projection period, the frequency of Vulnerable (A2-A4), and Vulnerable (A1)/Endangered(A2-A4) listings increased such that biomass declines of >= 30% were observed in 30% of simulations each year, and biomass declines of more than 50% were observed in ~15% of simulations (Fig 19C).

Conversely, the frequency at which the stock was assessed to be below biomass reference points dropped to very low levels following the initial stock recovery during the management period (Fig. 19D). There is thus an increased frequency of listings using IUCN Criterion A over the assessment of spawning stock status relative to reference points. This is reflected in the upper left quadrants of the confusion matrices (Fig 20A & 20B), with higher percentages of simulations triggering the listing criteria when the stock is assessed to be above reference points. This effect would be more pronounced if the matrixes for the estimates relative to the 30% decline threshold were shown.

Due to the low initial stock size, the probability that blackspot seabream fell below 10% spawning biomass was non-zero (4.9%) in the early years of the management period (Fig. 21).

3.5. Reference Point Comparison

The method for determining F_{MSY} and MSY $B_{trigger}$ reference points in the MSE simulations was not identical to that used by ICES, where these estimates are generally based on probabilistic outcomes. F_{MSY} is typically defined as being the F that results in the maximum of median catch from stock projections, and MSY $B_{trigger}$ being the 5th percentile of spawning biomass when fishing at F_{MSY} (ICES 2017). To compare the approach used to calculate the reference points in the MSE simulations with those obtained using the ICES Advice approach, estimates for both F_{MSY} and MSY $B_{trigger}$ were also obtained from the OMs by calculating these in the first assessment year of the MSE simulations (i.e. final year of the historical period). To solve for F_{MSY}, for each of the 1,000 simulations (for each stock), the OM numbers at age vector was multiplied by the assessment error for that year and then projected forwards with the known OM parameter values (but with assessment error in unfished stock size) for 100 years 100 times under a given F (with annual variation in recruitment) to obtain a set of time series of yield estimates. The median catch over the final 50 years of these 100 projections was used to calculate the estimate of catch resulting from the given F, and a golden search over F performed to maximize this median yield to determine the value for F_{MSY}. Estimates for MSY $B_{trigger}$ for each simulation were then derived from the 100 stock projections under F_{MSY} using the 5th percentile of spawning stock biomass over the same time period used to calculate F_{MSY}.

The distribution of estimates for the reference points using this alternative approach is compared to those used in the MSE simulations in Figures 24 and 25. For bluefin tuna, blue shark, and Baltic cod, stochastic estimates of F_{MSY} and MSY $B_{trigger}$ were lower than the deterministic estimates used in the simulations (Figs 24-25), although for blue shark, the distributions for the estimates of MSY $B_{trigger}$ from both approaches overlapped (Fig 25). The deterministic reference points for blackspot seabream were within the distribution of estimates obtained via the ICES approach. Aside from variability associated with assessment error and recruitment variability, the differences in reference points could be largely attributed to the mis-specification of the rate of natural mortality in the deterministic reference point method to the OM value (a fixed M at age was used in the calculations). The stochastic ICES advice reference point method was also run but with the stock projections using the same fixed M at age vector as in the approach used in the MSE rather than the true age-varying OM values. Using this fixed M largely corrected the difference in the two approaches, for bluefin tuna and Baltic cod for F_{MSY} (Fig. 24, "s_fixedM"), and for Baltic cod for MSY $B_{trigger}$ (Fig. 25, "s_fixed_M"). Blackspot seabream showed little sensitivity because the OM values for natural mortality did not vary over age. Fixing M as constant over age, also accounted for some of the differences in estimates for blue shark (both reference points), and for tuna MSY $B_{trigger}$.

4. Discussion

The MSE simulations for the four stocks demonstrate the utility of assessing stocks against reference points that reflect stock status rather than just stock trend. The simulations also show that management using a control rule that explicitly reduces fishing pressure on a stock when it is assessed to be at low levels is successful at maintaining populations above critical biomass thresholds – even when considering process error, estimation error, and model error in the stochastic projections of stock development – extinction probability is hard to define in these types of population dynamics models, yet in all four species case studies, stocks either never or very rarely fell below 10% of unfished spawning biomass.

- The simulations showed a degree of non-correspondence between assessments that triggered IUCN listings under criterion A and that also assessed stocks to be below biomass reference points – for those simulated assessments when stocks were assessed to be below reference points, the percentage of assessments where listing thresholds were met was low. However, in many cases listings against categories of conservation concern under IUCN criteria A may have likely already occurred based on estimates of historical stock dynamics, particularly for these case studies, which largely featured depleted stocks, regardless of data quantity. This suggests however that listings under IUCN criteria A are unlikely to be made during stock recovery (frequency of assessments resulting in listings only increased towards the end of the management period when stocks were fluctuating around biomass 'targets'), even if stocks still remain at low levels relative to traditional fisheries management reference points. A consequence of this is that if management of a fishery were to be based solely on the outcome of IUCN listing assessments, then the frequency at which those assessments were conducted in the future would be a crucial determinant of stock development and yield outcome.

- Stocks beginning management at higher stock sizes showed better correspondence of reference point status assessment and listing decisions: blue shark had the 'best' looking confusion matrix. A second scenario was run for blue shark where the initial stock size prior to management was 40% of unfished spawning biomass (Figs 21 & 22). In this scenario, listing frequencies and status relative to biomass reference points were even more comparable, particularly with regard to status relative to B_{lim} (Fig 22B), where 70% of simulated assessments estimating stock to be below B_{lim} also being associated with an estimated biomass decline of >=50%. In this case, a low stock size is much more frequently associated with a biomass decline (through the effects of recruitment variability and imprecision and bias in catch advice).
- The combined IUCN listing assessments for the Baltic cod simulations that summed biomass over both eastern and western stocks to determine the magnitude of biomass change showed the potential problem with this approach, where one of the stocks (most likely the smaller one) could be below limit reference points yet a listing would not occur until both stocks showed biomass decline. This effect would possibly be exacerbated when stock trajectories and dynamics differed between stocks.
- Populations were assumed to be unexploited with equilibrium age structure in the initial year of the simulations – for these stocks that have been exploited for many years this assumption means that population trajectory during the initial part of the historical period likely does not properly reflect that observed, as the fishing mortality rates in the operating model determined by the actual historical catches during this period will be higher. The effect of this in the simulations here is likely to inflate the frequency of listings under IUCN categories with criterion A made during the historical period before implementation of the management strategy.
- Single stock dynamics and single fisheries were modelled, with removals only coming from the fishing fleet under effect of the management action. This is unlikely to realistically mimic the case for mixed populations – such as for Bluefin tuna where there is mixing between western and eastern stocks and uncertainty in the assignment of fishery removals to population (and indeed the ability to estimate population status effectively). This could have at least two effects: first, the masking effect observed in the combined Baltic cod analyses against IUCN thresholds may be more prevalent, and

second, there will be increased bias and imprecision in the estimates of stock size and reference point determination. Without explicitly modelling additional populations within the same simulations, these effects could be approximated in the current framework by including implementation error and/or additional bias in the pseudo-assessments.

- The pseudo-assessments, while computationally simple, do not capture the full structure and characteristics of a stock assessment – data quality and availability is simply reflected in the magnitude of bias and imprecision of the assessment estimate. Nevertheless, this approach has been used successfully to approximate the behaviour of stock assessment process. This modelling decision allowed for the evaluation of a larger number of species/stock scenarios, while still retaining the structure necessary for a full MSE feedback loop scenario.
- Similar to the pseudo-assessment, the control rule and reference point determination used in the scenarios was an approximation to that implemented by ICES. The pseudo-assessments did not freely estimate reference points in each assessment instance, meaning estimated reference points were closer to the true operating model values than they would be were these estimated within a stock assessment process. ICES reference points are (when possible) based on probabilistic outcomes. The comparison of these approaches suggested that lower values for reference points were obtained under the probabilistic approach but that the difference between estimates could largely be attributed to mis-specification of natural mortality. Consequences of decreased values for reference points might suggest that the ICES Advice rule would be less sensitive to changes in biomass (because F would not be reduced until lower stock size), but that the intensity of fishing overall would be lower. This might translate to higher overall stock sizes were this reference point estimation approach included within the MSE simulations. The analyses conducted here probably define less conservative (stock-wise) reference points than would be obtained through the full computationally intensive reference point determination procedures.
- The simulations showed very low probabilities of stock collapse, suggesting that the control rule performs well at buffering the effects of population decline. However, stock dynamics did not include all effects and so true determination of stock dynamics at low stock sizes is hard to make inference about. In theory, the consequences of small population size factors (e.g. recruitment failure) could be included in the current analyses by modelling the probability of deleterious events (like a recruitment failure) as a function of stock size. Conditioning these scenarios on appropriate values for the relevant parameters might be difficult.
- In MSE it is important that the OM dynamics adequately represent the population dynamics of the modelled stock. This may not have been the case for the blackspot seabream results - stocks of blackspot seabream have been presumably at low levels for some time with very low levels of fishing. Given the estimated stock size at the start of the simulations, the control rule would have been conservative in catch levels – in the simulations the stock responded and recovered fairly quickly. While conditioning an operating model by fitting to available data is a preferred approach to ensuring dynamics are consistent with available data, in the absence of stock assessments an alternative approach could be to do rejection sampling on the

simulated historical stock trajectories to better match the characteristics of the observed stock behaviour. The current MSE framework used here offers a rapid approach for doing this.

- MSE is a powerful approach for comparing the relative performance and robustness of alternative management strategies. However, given that appropriate levels and dimensions of uncertainty are accounted for in stock dynamics and model structure, MSE offers an opportunity to obtain probability distributions for stock outcomes given particular management, that fully integrates both uncertainty in the biological dynamics but also uncertainty in the fishery system, including estimation of stock status, estimation and evaluation of reference points, management error, and implementation error resulting from the uncertainty in fishery responses to changes in management. This offers a more complete package of uncertainty associated with management of social-ecological systems than a traditional Population Viability Analysis, suggesting that MSE is an appropriate tool to obtain quantitative estimates of probabilities for stock outcomes and could be a useful tool for assessing fish populations under IUCN criterion E.

5. References

Bunnefeld, N., Hoshino, E. and Milner-Gulland, E.J., 2011. Management strategy evaluation: a powerful tool for conservation? Trends in ecology & evolution, 26(9), pp.441-447.

Fay, G. and G.N. Tuck. (eds.) 2011. Development of a multi-gear spatially explicit assessment and management strategy evaluation for the Macquarie Island Patagonian toothfish fishery. Australian Fisheries Management Authority and CSIRO Marine and Atmospheric Research, Hobart. 181p.

Fay, G., Punt, A.E. and Smith, A.D., 2011. Impacts of spatial uncertainty on performance of age structure-based harvest strategies for blue eye trevalla (Hyperoglyphe antarctica). Fisheries Research, 110(3), pp.391-407.

ICCAT, 2015. Report Of The 2015 Iccat Blue Shark Stock Assessment Session. (Oceanário de Lisboa, Lisbon, Portugal - July 27 to 31, 2015). 116p.

ICCAT, 2017. Report of the standing committee on research and statistics (SCRS). Madrid, Spain, October 2017. 465p.

ICES. 2015. Report of the Benchmark Workshop on Baltic Cod Stocks (WKBALTCOD), 2–6 March 2015, Rostock, Germany. ICES CM 2015/ACOM:35. 172 pp.

ICES, 2017. ICES fisheries management reference points for category 1 and 2 stocks. DOI: 10.17895/ices.pub.3036

ICES. 2018. ICES Advice on fishing opportunities, catch, and effort. Baltic Sea Ecoregion. Cod (Gadus morhua) in subdivisions 24–32, eastern Baltic stock (eastern Baltic Sea). https://doi.org/10.17895/ices.pub.4378

Klaer, N.L., Wayte, S.E. and Fay, G., 2012. An evaluation of the performance of a harvest strategy that uses an average-length-based assessment method. Fisheries Research, 134, pp.42-51.

Little, L.R., Parslow, J., Fay, G., Grafton, R.Q., Smith, A.D., Punt, A.E. and Tuck, G.N., 2014. Environmental derivatives, risk analysis, and conservation management. Conservation Letters, 7(3), pp.196-207.

Lorance, P., 2010. History and dynamics of the overexploitation of the blackspot sea bream (Pagellus bogaraveo) in the Bay of Biscay. ICES Journal of Marine Science, 68(2), pp.290-301.

Punt, A.E., Dorn, M.W. and Haltuch, M.A., 2008. Evaluation of threshold management strategies for groundfish off the US West Coast. Fisheries Research, 94(3), pp.251-266.

Wiedenmann, J., Wilberg, M.J., Sylvia, A. and Miller, T.J., 2015. Autocorrelated error in stock assessment estimates: implications for management strategy evaluation. Fisheries research, 172, pp.325-334.

Table 1: Operating model specifications

	Bluefin tuna	Blue shark	Baltic cod	Blackspot seabream
Historical period	1950–2015	1968–2013	1950–2016	1963–2010
Length range (cm)	20–350	30–450	5–125	5–85
Length bin width	10	10	5	5
Maximum Age	16	16	7	30
Natural Mortality (age 0-max.)	0.41, 0.41, 0.32, 0.26, 0.22, 0.19, 0.17, 0.15, 0.14, 0.13, 0.12, 0.12, 0.11, 0.11, 0.11, 0.11, 0.11, 0.11, 0.11, 0.11, 0.10	0.38, 0.30, 0.27, 0.25, 0.24, 0.23, 0.22, 0.22, 0.21, 0.21, 0.21, 0.21, 0.20, 0.20, 0.20, 0.20, 0.20	0.8, 0.242, 0.2, 0.2, 0.2, 0.2, 0.2, 0.2	0.2 (all ages)
Length at age				
VB $L_{infinity}$	319	310, 282	89.3	51.4
VB k	0.093	0.12, 0.18	0.187	0.137
VB t_0	-0.97	-1.77, -1.35	-0.1515	-0.97
CV length at age (age 0 & max)	0.1, 0.1	0.15, 0.12	0.1, 0.1	0.1, 0.1
Weight-length a	0.0000177054	3.18e-06	6.0e-06	1.4e-05
Weight-length b	3.001251847	3.1313	3.1503	3.017
Length at maturity (cm)	195 (Fig 1C)	195 (Fig 2C, max. fecundity = 39 pups)	Fig 3C	30 (50% maturity, Fig 4C)
Stock-Recruitment				
Final depletion	0.15	0.25 (0.4)	0.125	0.1
Steepness, h	0.55	0.7	0.8	0.65
Recruitment variance	0.73	0.4	0.6	0.4
Source for OM parameter values	ICCAT (2017)	ICCAT (2015)	ICES (2018), ICES (2015)	Lorance (2010)

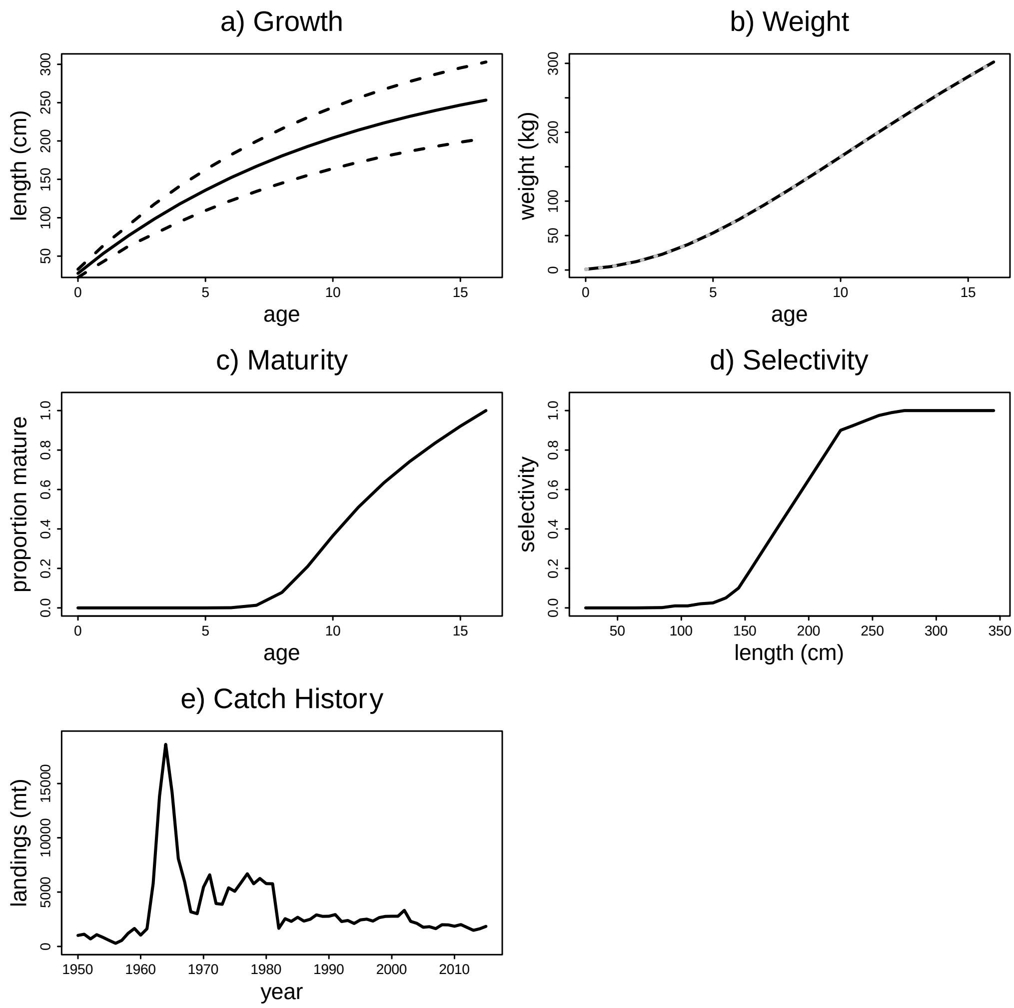


Figure 7: Operating model specifications for western Bluefin tuna showing a) mean (solid line) and standard deviation (dashed line) of length at age, b) weight at age (solid line females, dashed line males), c) maturity at age, d) selectivity at length, and e) time series of catches.

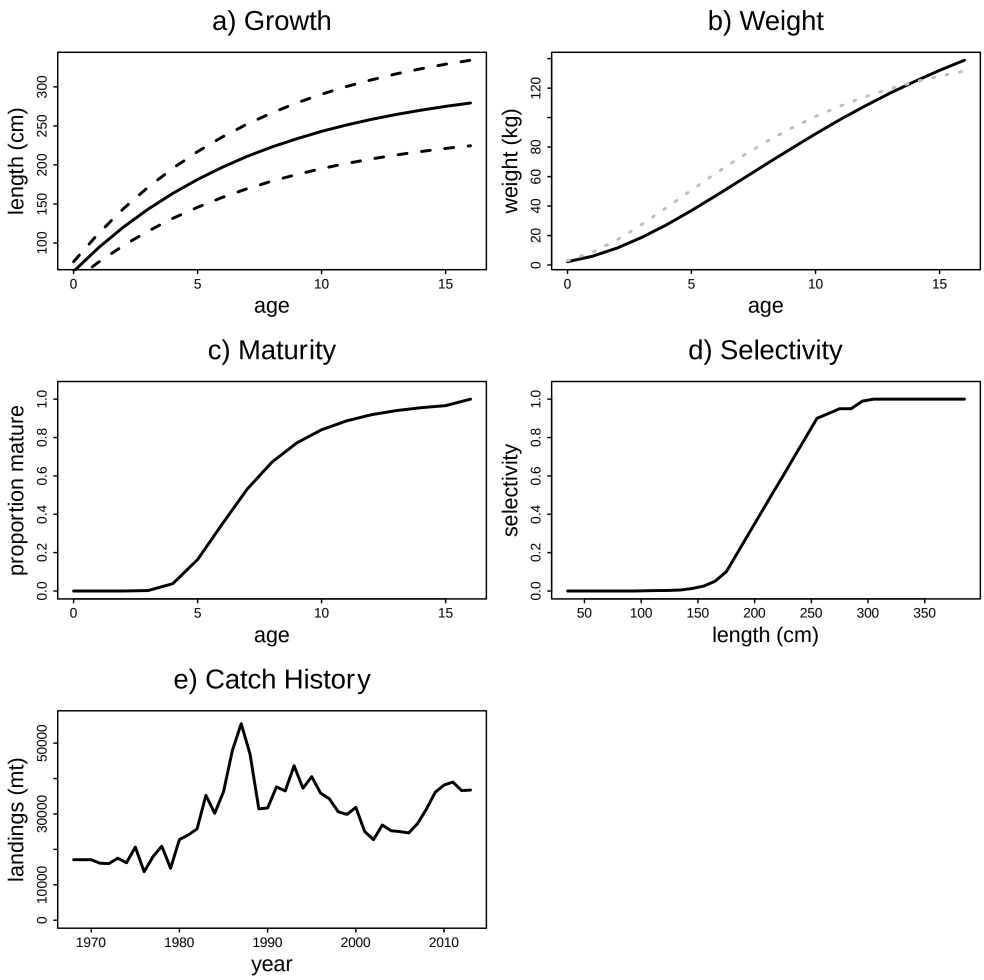


Figure 8: Operating model specifications for Atlantic blue shark showing a) mean (solid line) and standard deviation (dashed line) of length at age, b) weight at age (solid line females, dashed line males), c) maturity at age, d) selectivity at length, and e) time series of catches.

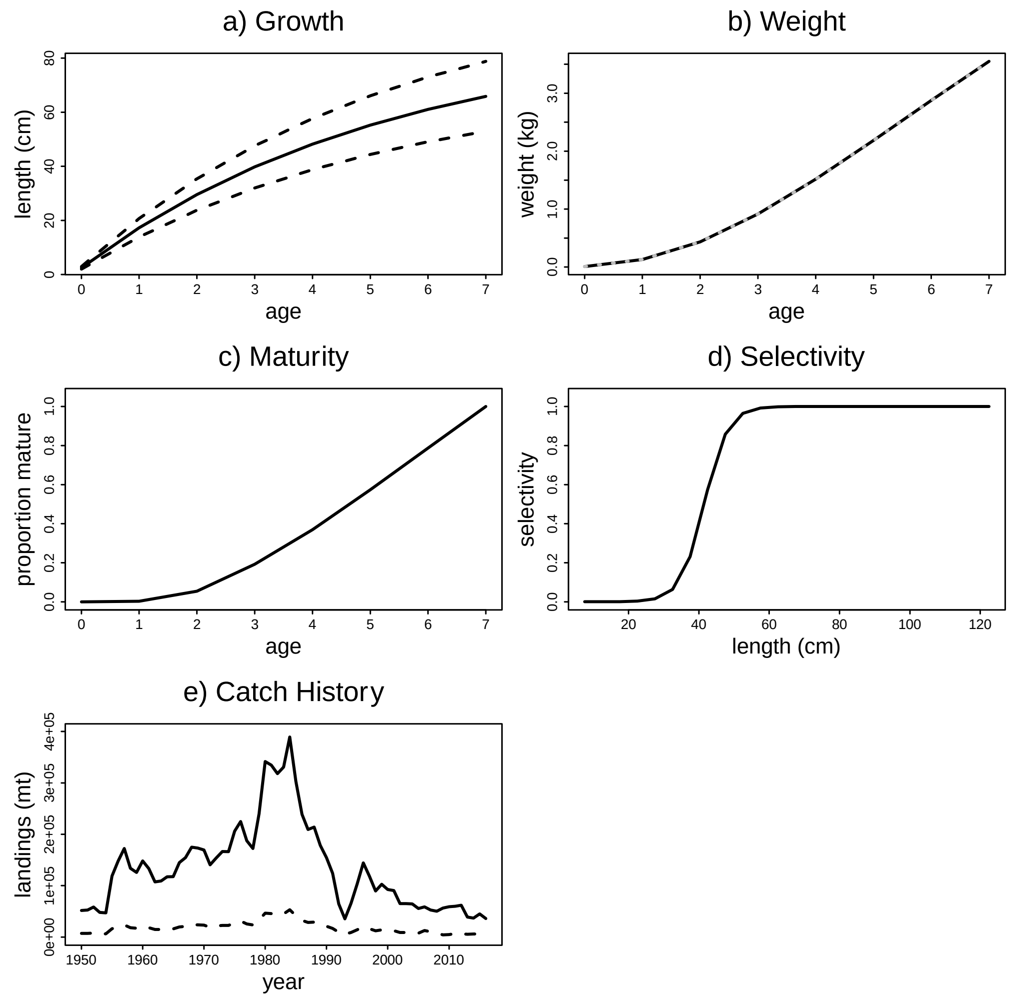


Figure 9: Operating model specifications for Baltic cod showing a) mean (solid line) and standard deviation (dashed line) of length at age, b) weight at age (solid line females, dashed line males), c) maturity at age, d) selectivity at length, and e) time series of catches (eastern, solid line; western, dashed line).

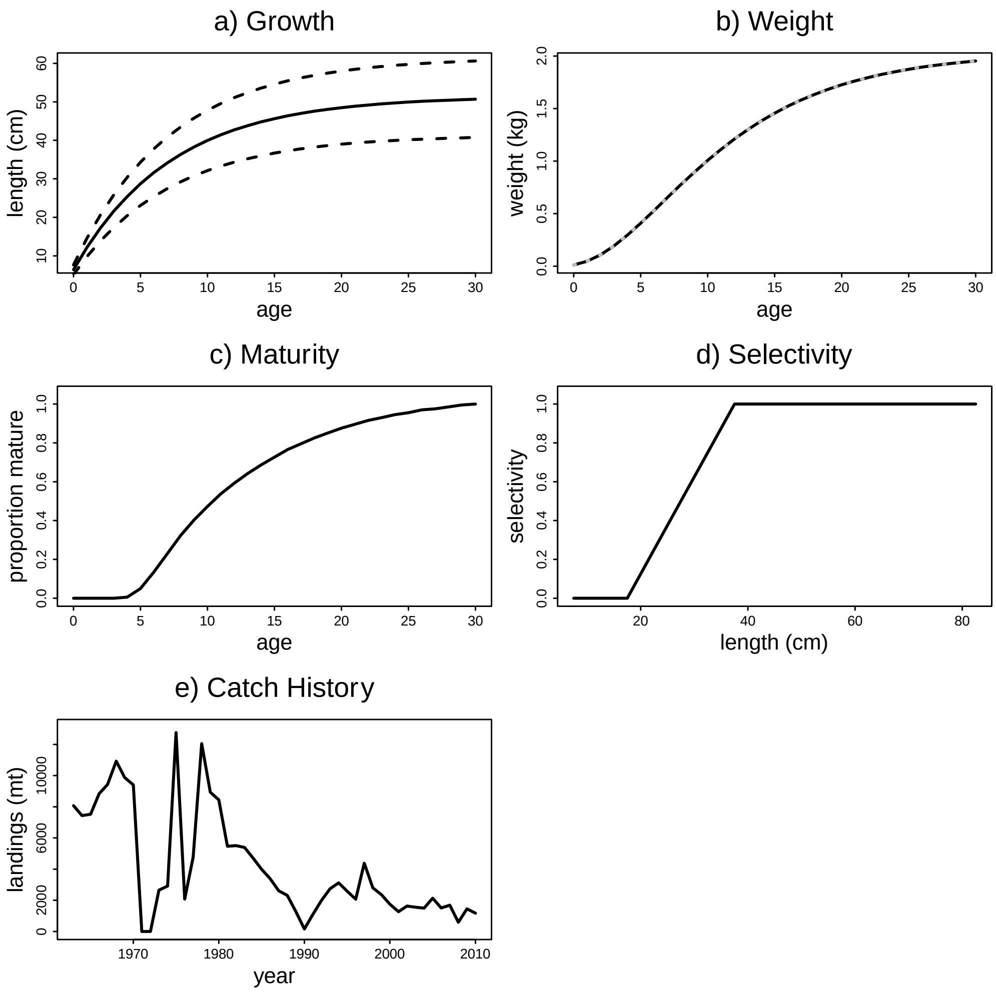


Figure 10: Operating model specifications for blackspot seabream showing a) mean (solid line) and standard deviation (dashed line) of length at age, b) weight at age (solid line females, dashed line males), c) maturity at age, d) selectivity at length, and e) time series of catches.

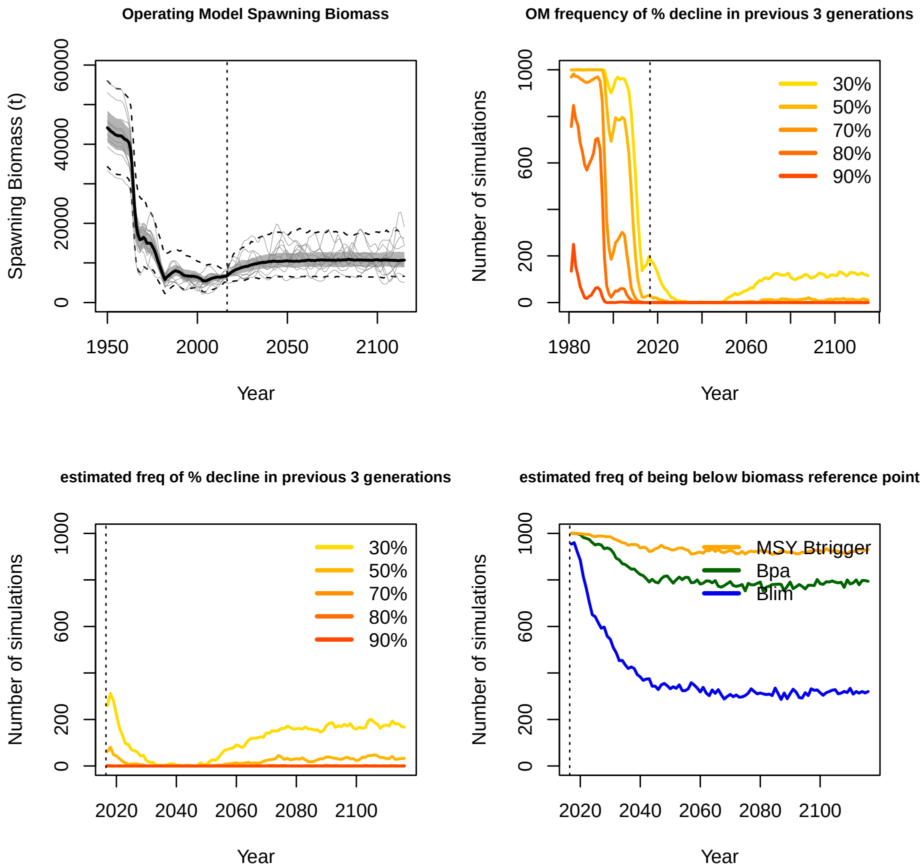


Figure 11: Time series of outputs for western Atlantic Bluefin tuna simulations. (a) summary of spawning biomass trajectory, with median (dark line), interquartile range (grey shading), and 95% interval (dashed lines). Example individual simulation trajectories shown in grey lines. (b) Number of simulations in which change in operating model spawning biomass over three previous generations exceeds thresholds for IUCN categories, (c) Number of simulations in which the estimated change in \spawning biomass over three previous generations exceeds thresholds for IUCN categories, (d) number of simulations in which spawning biomass is estimated to be below ICES biomass reference points.

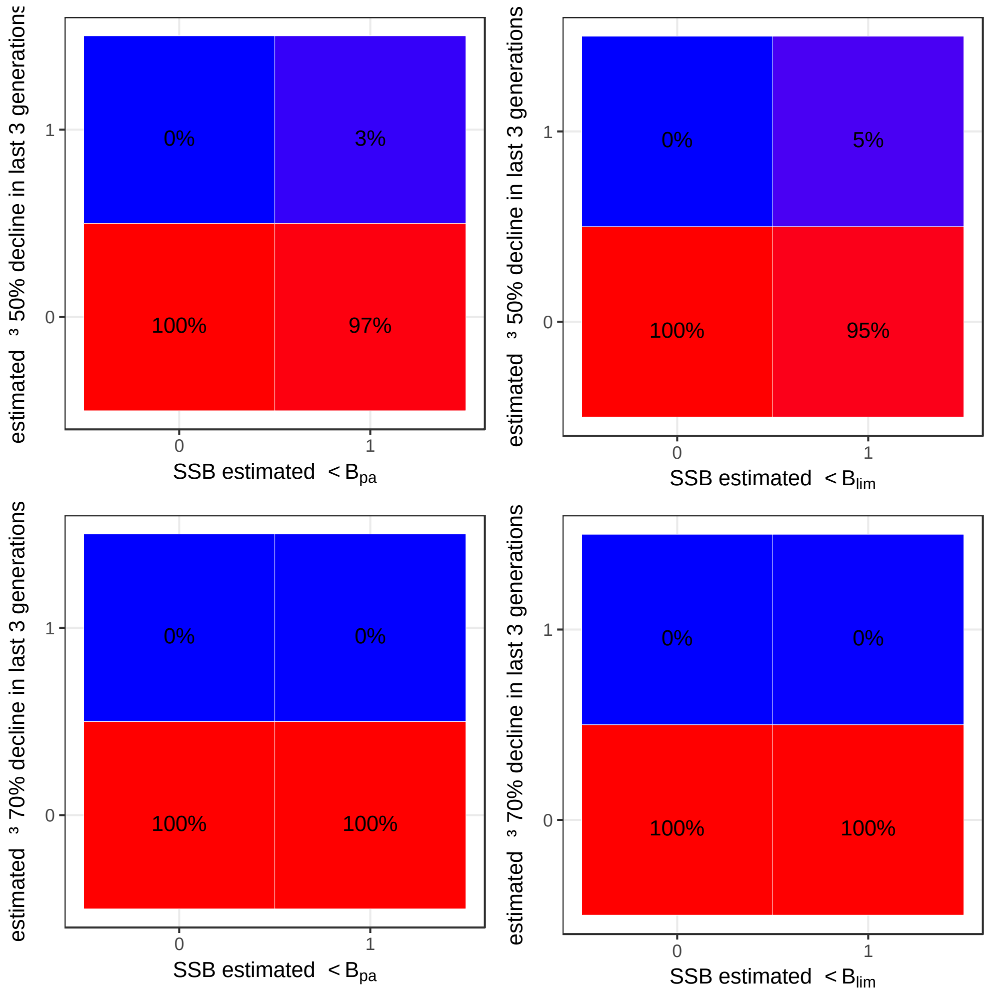


Figure 12: Confusion matrices for western Atlantic Bluefin tuna, with percentage of simulated assessments (over both projection years and simulations) where biomass change is estimated to trigger thresholds for IUCN criterion A (1) or not (0) given that spawning biomass is estimated to be below either B_{pa} or B_{lim} (1) or not (0).

Figure 13: Probability of operating model spawning biomass being below 10%, 5%, and 1% of unfished spawning biomass for western Atlantic Bluefin tuna.

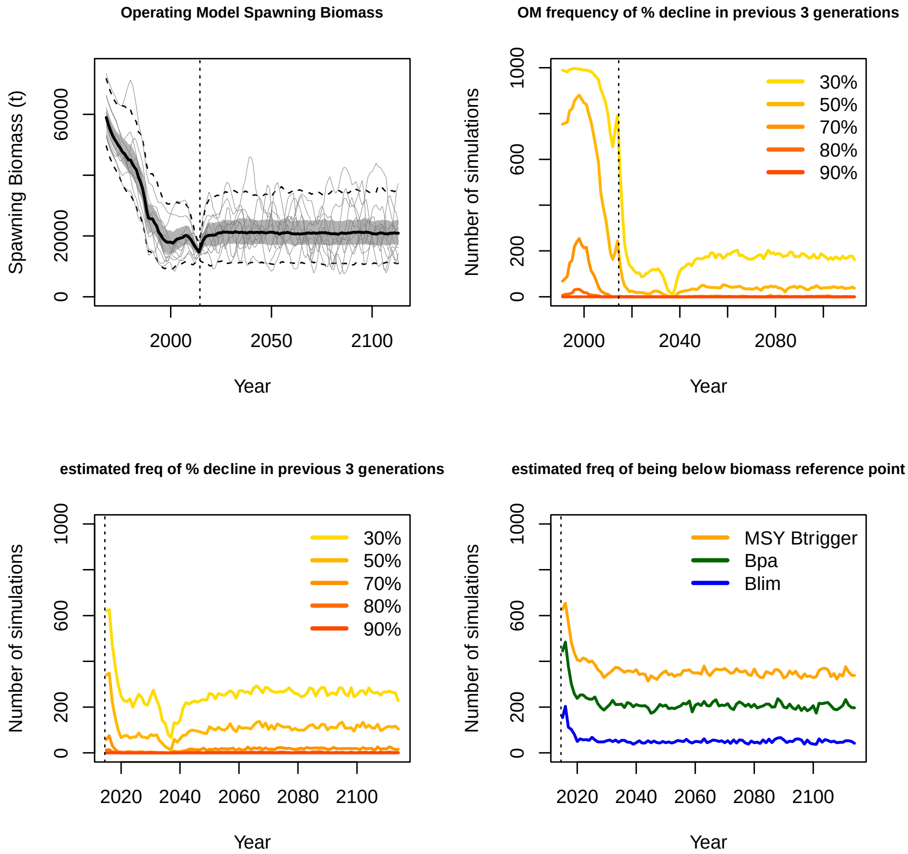


Figure 14: Time series of outputs for blue shark simulations. (a) summary of spawning biomass trajectory, with median (dark line), interquartile range (grey shading), and 95% interval (dashed lines). Example individual simulation trajectories shown in grey lines. (b) Number of simulations in which change in operating model spawning biomass over three previous generations exceeds thresholds for IUCN categories, (c) Number of simulations in which the estimated change in \spawning biomass over three previous generations exceeds thresholds for IUCN categories, (d) number of simulations in which spawning biomass is estimated to be below ICES biomass reference points.

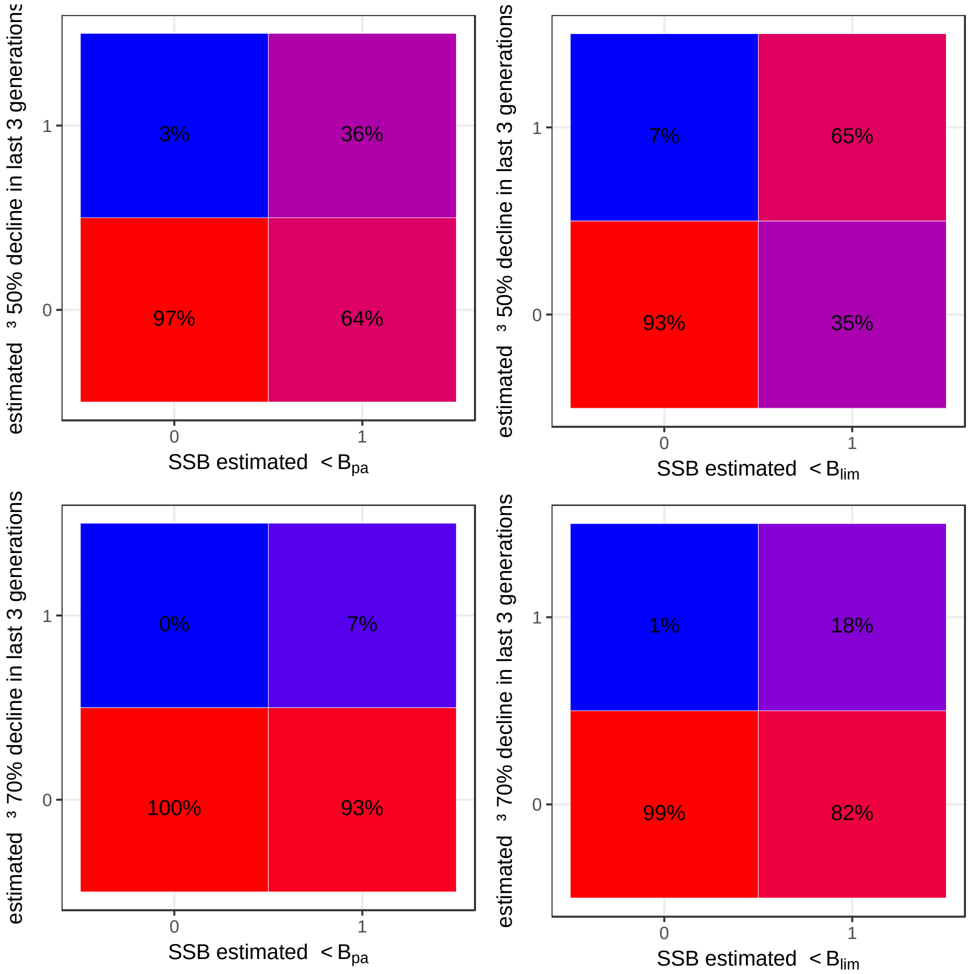


Figure 15: Confusion matrices for blue shark, with percentage of simulated assessments (over both projection years and simulations) where biomass change is estimated to trigger thresholds for IUCN criterion A (1) or not (0) given that spawning biomass is estimated to be below either B_{pa} or B_{lim} (1) or not (0).

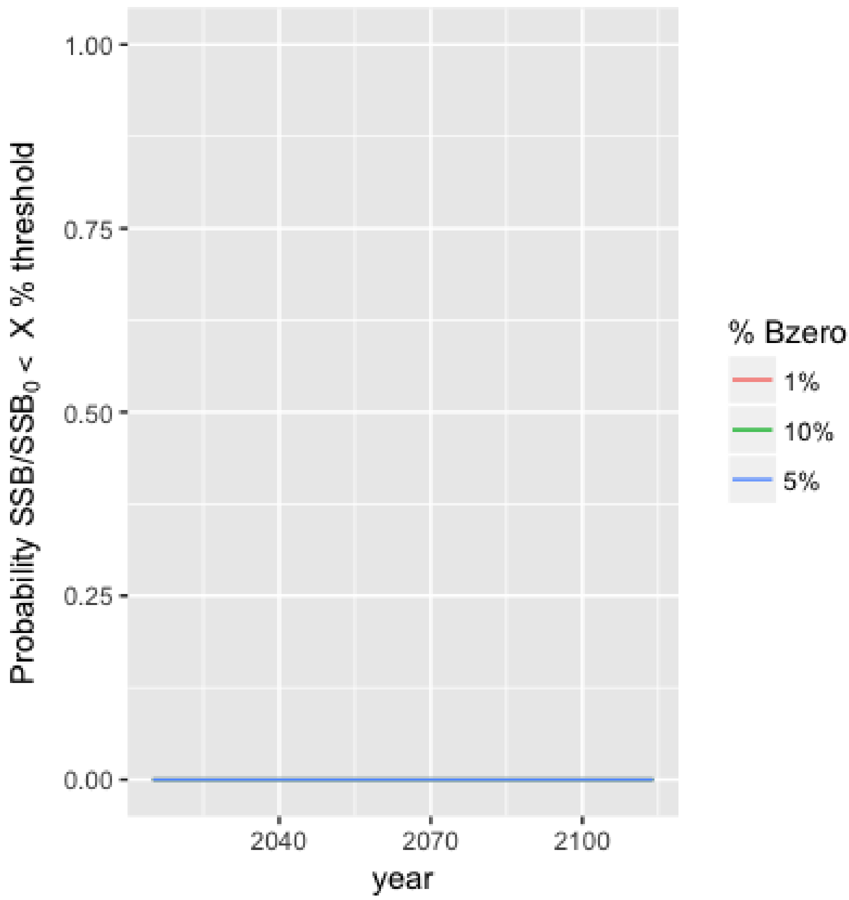


Figure 16: Probability of operating model spawning biomass being below 10%, 5%, and 1% of unfished spawning biomass for blue shark.

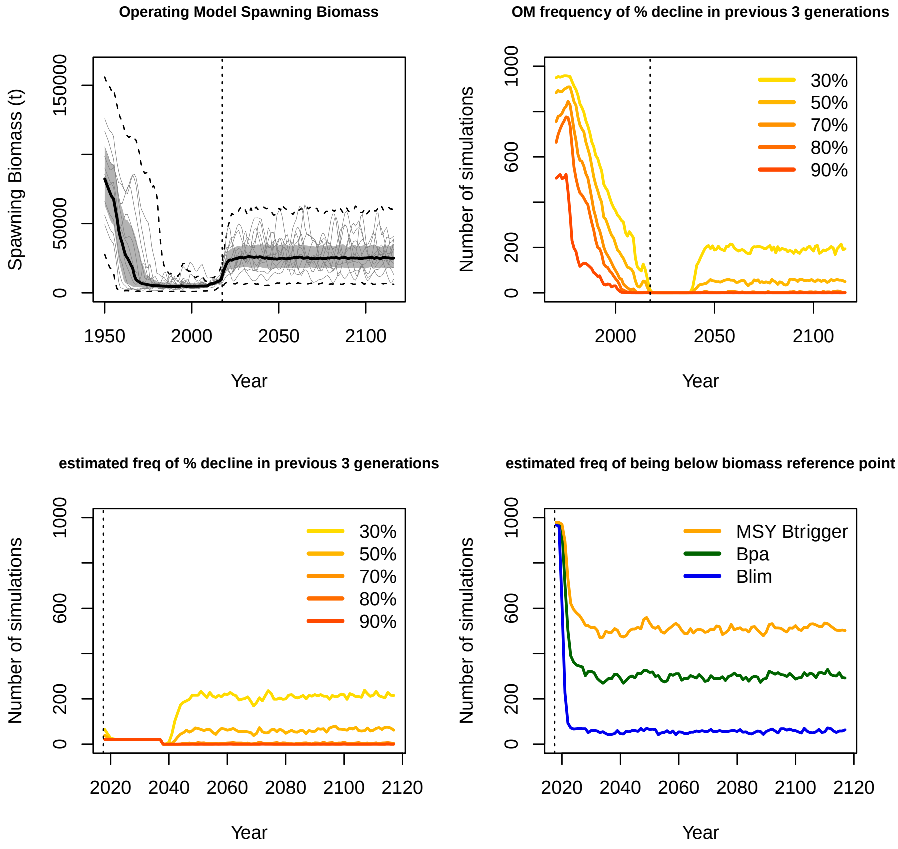


Figure 17: Time series of outputs for western Baltic cod simulations. (a) summary of spawning biomass trajectory, with median (dark line), interquartile range (grey shading), and 95% interval (dashed lines). Example individual simulation trajectories shown in grey lines. (b) Number of simulations in which change in operating model spawning biomass over three previous generations exceeds thresholds for IUCN categories, (c) Number of simulations in which the estimated change in \spawning biomass over three previous generations exceeds thresholds for IUCN categories, (d) number of simulations in which spawning biomass is estimated to be below ICES biomass reference points.

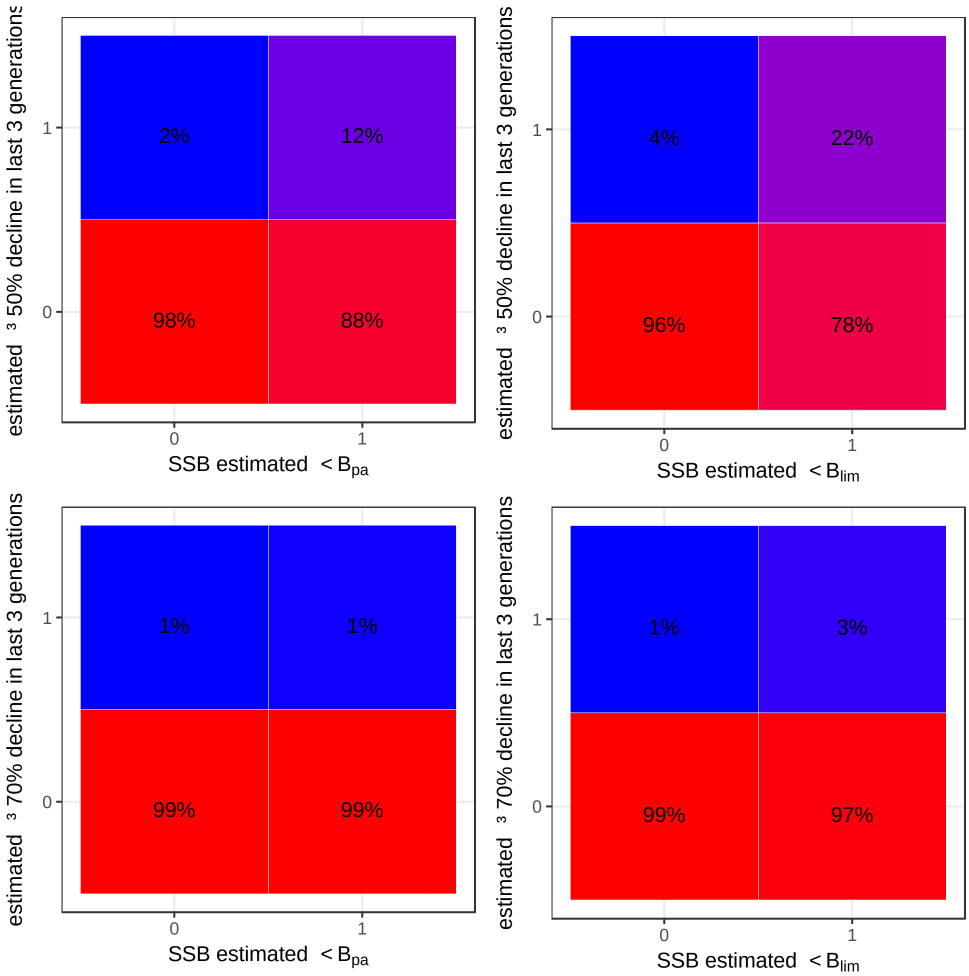


Figure 18: Confusion matrices for western Baltic cod, with percentage of simulated assessments (over both projection years and simulations) where biomass change is estimated to trigger thresholds for IUCN criterion A (1) or not (0) given that spawning biomass is estimated to be below either B_{pa} or B_{lim} (1) or not (0).

Figure 19: Probability of operating model spawning biomass being below 10%, 5%, and 1% of unfished spawning biomass for western Baltic cod.

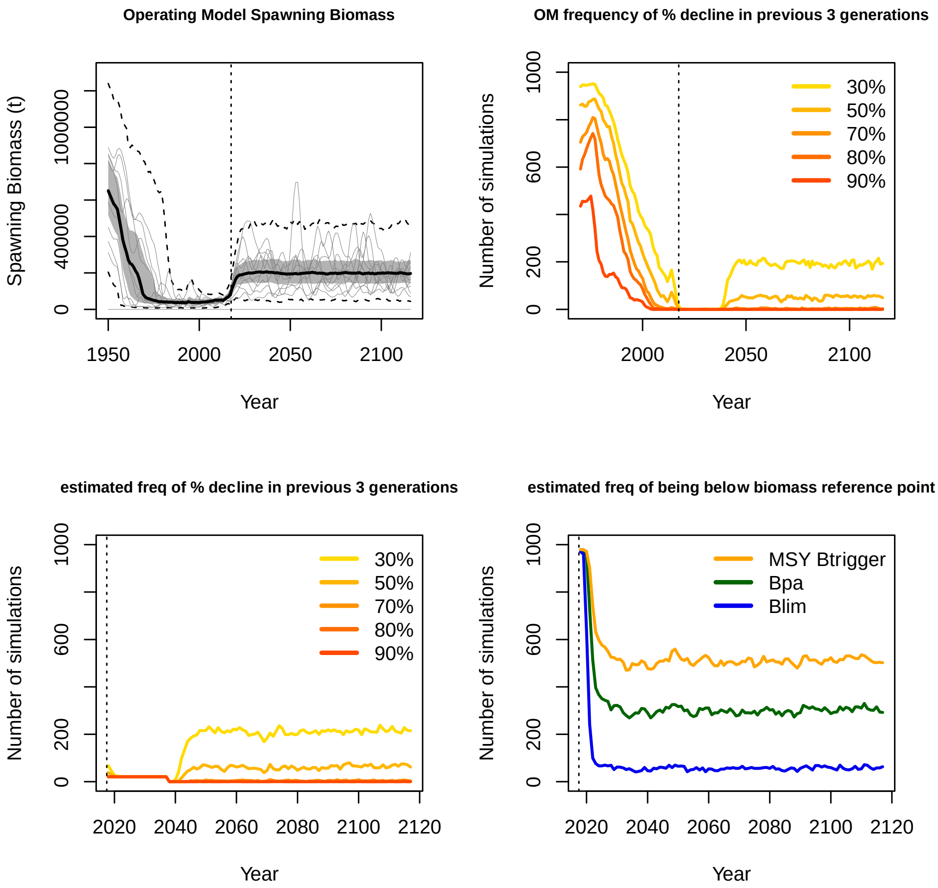


Figure 20: Time series of outputs for eastern Baltic cod simulations. (a) summary of spawning biomass trajectory, with median (dark line), interquartile range (grey shading), and 95% interval (dashed lines). Example individual simulation trajectories shown in grey lines. (b) Number of simulations in which change in operating model spawning biomass over three previous generations exceeds thresholds for IUCN categories, (c) Number of simulations in which the estimated change in \spawning biomass over three previous generations exceeds thresholds for IUCN categories, (d) number of simulations in which spawning biomass is estimated to be below ICES biomass reference points.

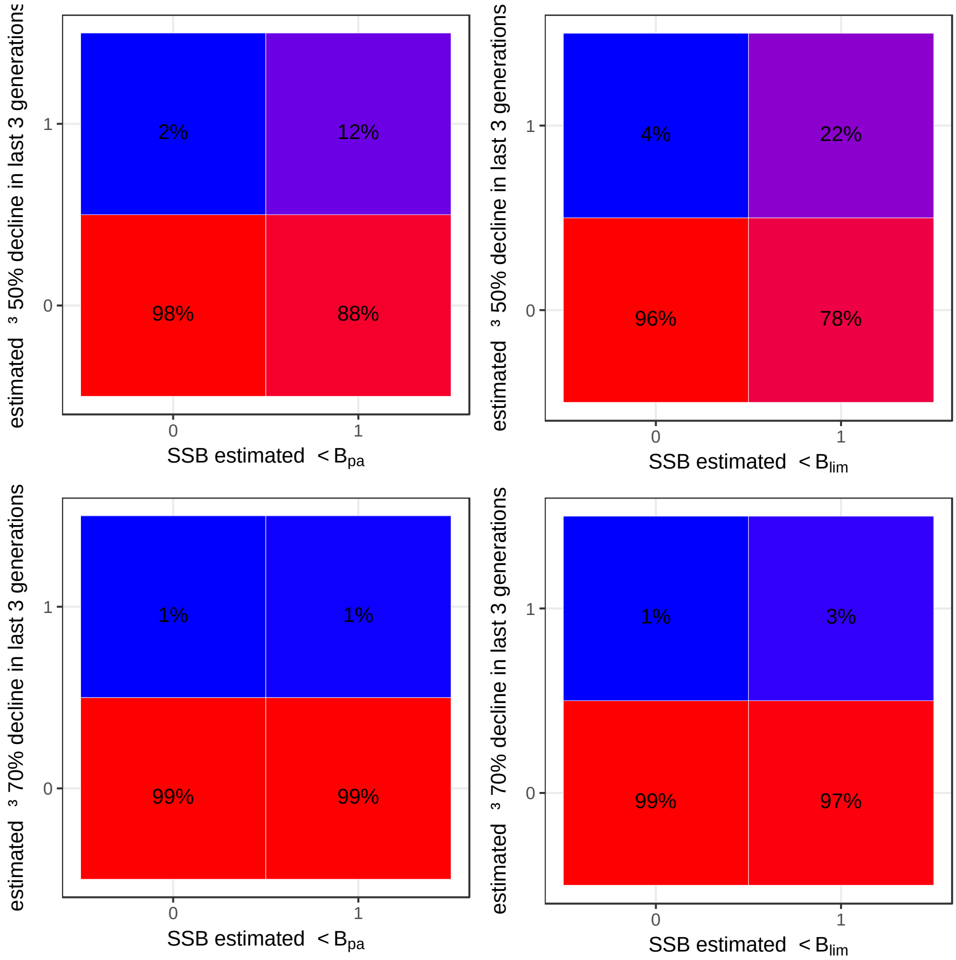


Figure 21: Confusion matrices for eastern Baltic cod, with percentage of simulated assessments (over both projection years and simulations) where biomass change is estimated to trigger thresholds for IUCN criterion A (1) or not (0) given that spawning biomass is estimated to be below either B_{pa} or B_{lim} (1) or not (0).

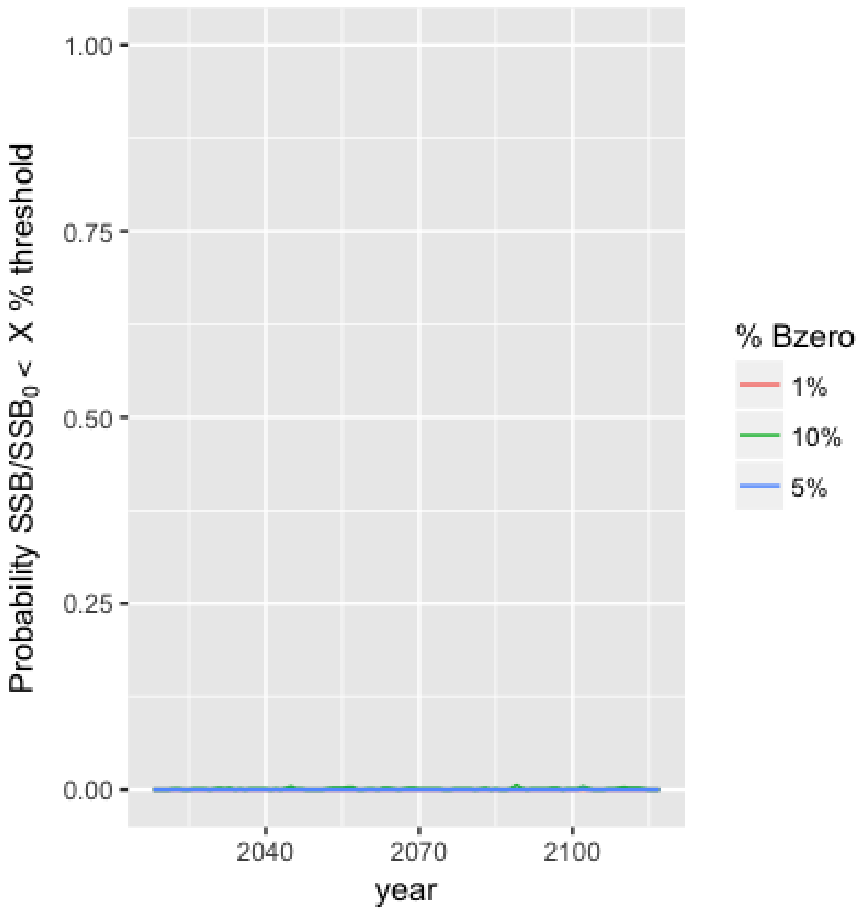


Figure 22: Probability of operating model spawning biomass being below 10%, 5%, and 1% of unfished spawning biomass for eastern Baltic cod.

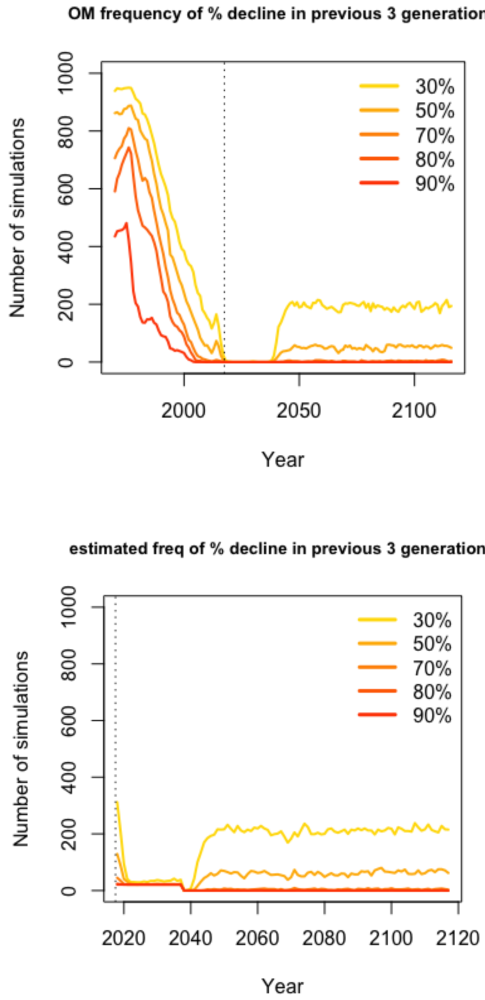


Figure 23: Time series of outputs for Baltic cod simulations when separate west and east stocks exist but metrics of biomass change are based on aggregate level (summed over both stocks). (a) Number of simulations in which change in operating model spawning biomass over three previous generations exceeds thresholds for IUCN categories, (b) Number of simulations in which the estimated change in \spawning biomass over three previous generations exceeds thresholds for IUCN categories.

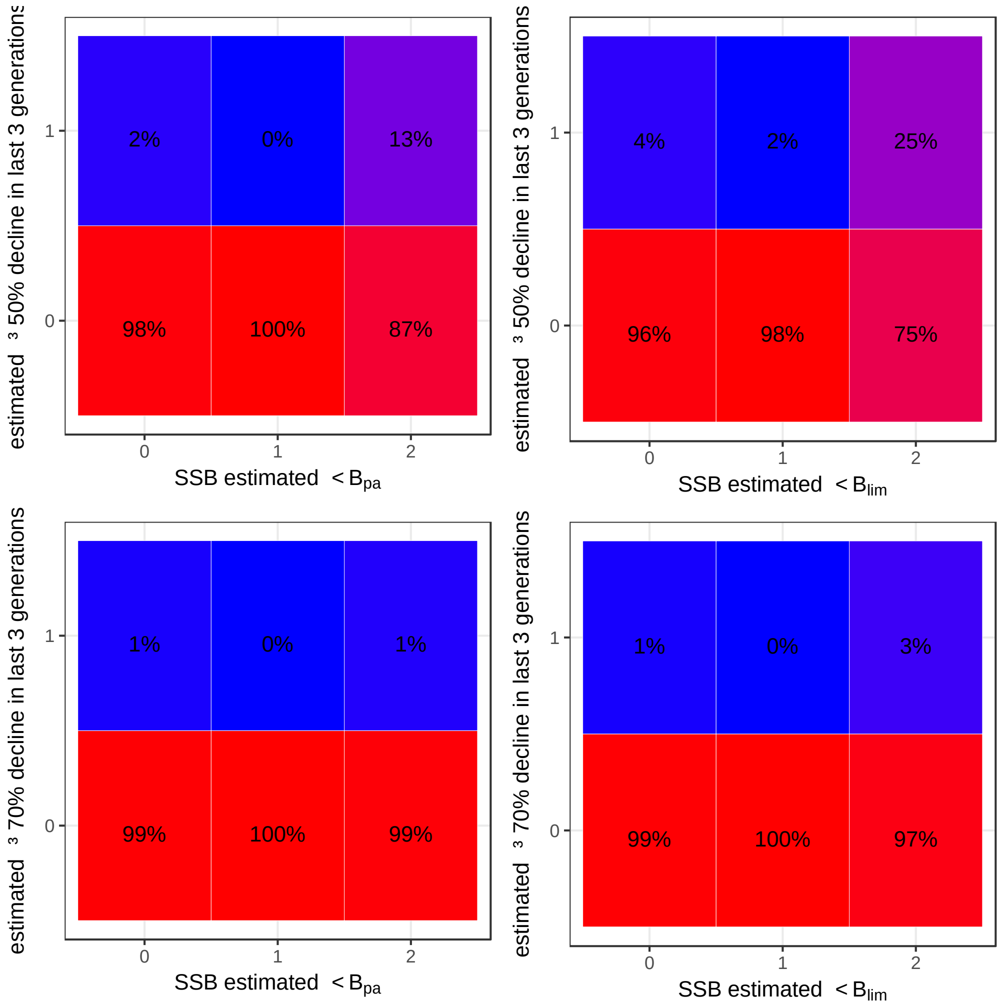


Figure 24: Confusion matrices for Baltic cod, with percentage of simulated assessments (over both projection years and simulations) where aggregated (western + eastern) biomass change is estimated to trigger thresholds for IUCN criterion A (1) or not (0) given that spawning biomass is estimated to be below either B_{pa} or B_{lim} in either one (1) or both (2) of the western and eastern stocks, or not (0).

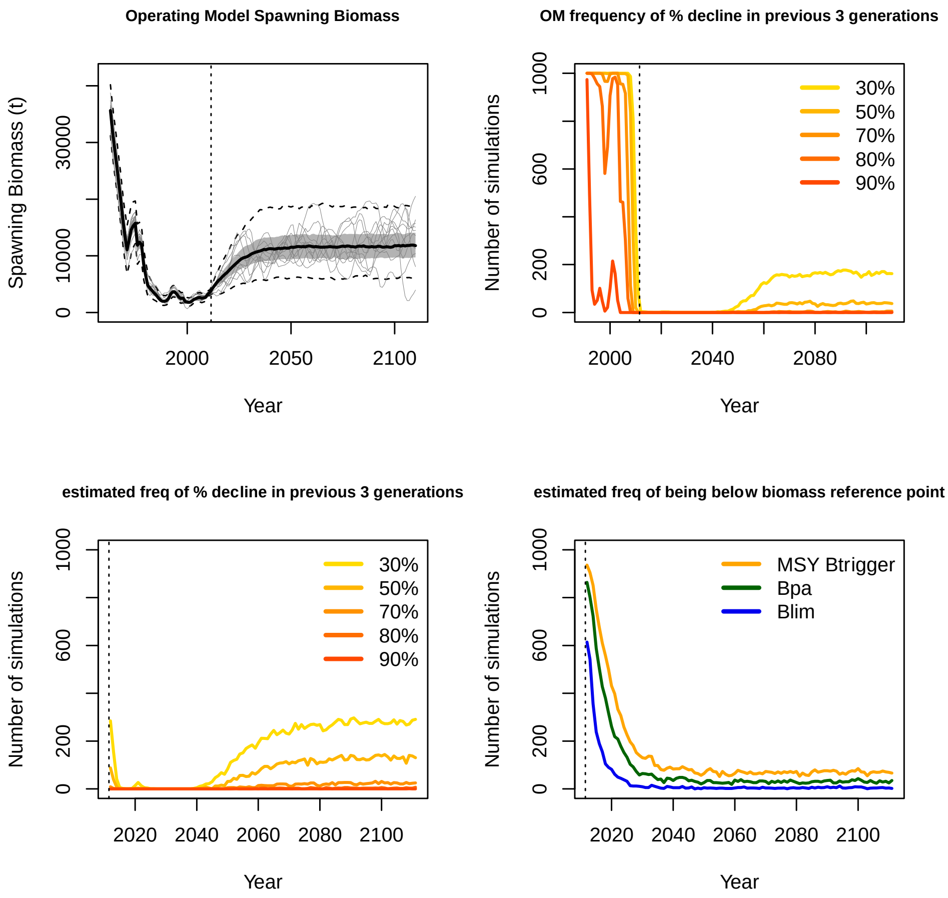


Figure 25: Time series of outputs for blackspot seabream simulations. (a) summary of spawning biomass trajectory, with median (dark line), interquartile range (grey shading), and 95% interval (dashed lines). Example individual simulation trajectories shown in grey lines. (b) Number of simulations in which change in operating model spawning biomass over three previous generations exceeds thresholds for IUCN categories, (c) Number of simulations in which the estimated change in \spawning biomass over three previous generations exceeds thresholds for IUCN categories, (d) number of simulations in which spawning biomass is estimated to be below ICES biomass reference points.

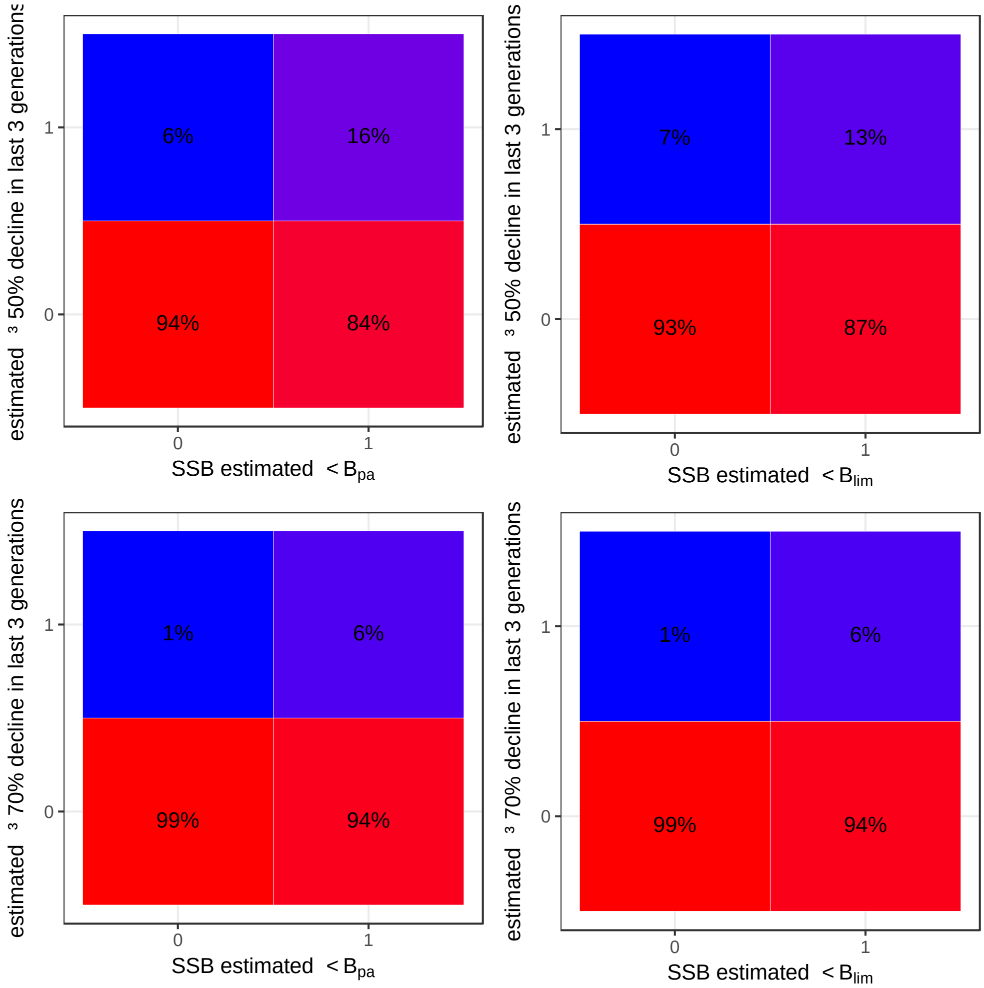


Figure 26: Confusion matrices for blackspot seabream, with percentage of simulated assessments (over both projection years and simulations) where biomass change is estimated to trigger thresholds for IUCN criterion A (1) or not (0) given that spawning biomass is estimated to be below either B_{pa} or B_{lim} (1) or not (0).

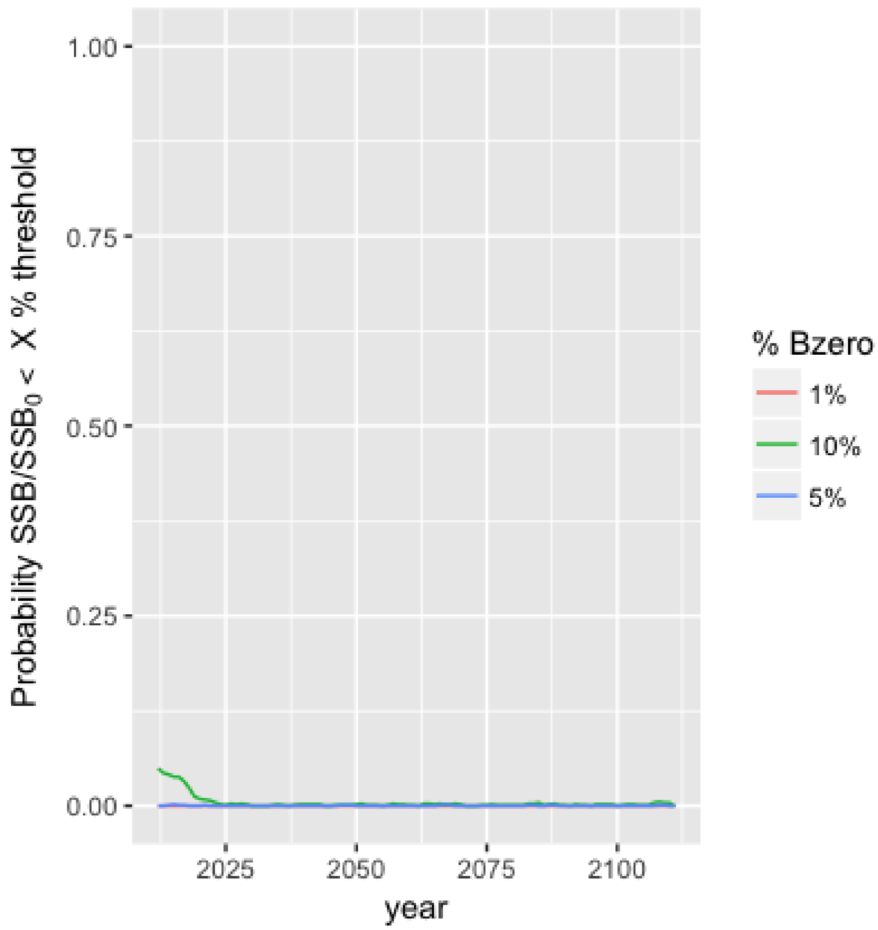


Figure 27: Probability of operating model spawning biomass being below 10%, 5%, and 1% of unfished spawning biomass for blackspot seabream.

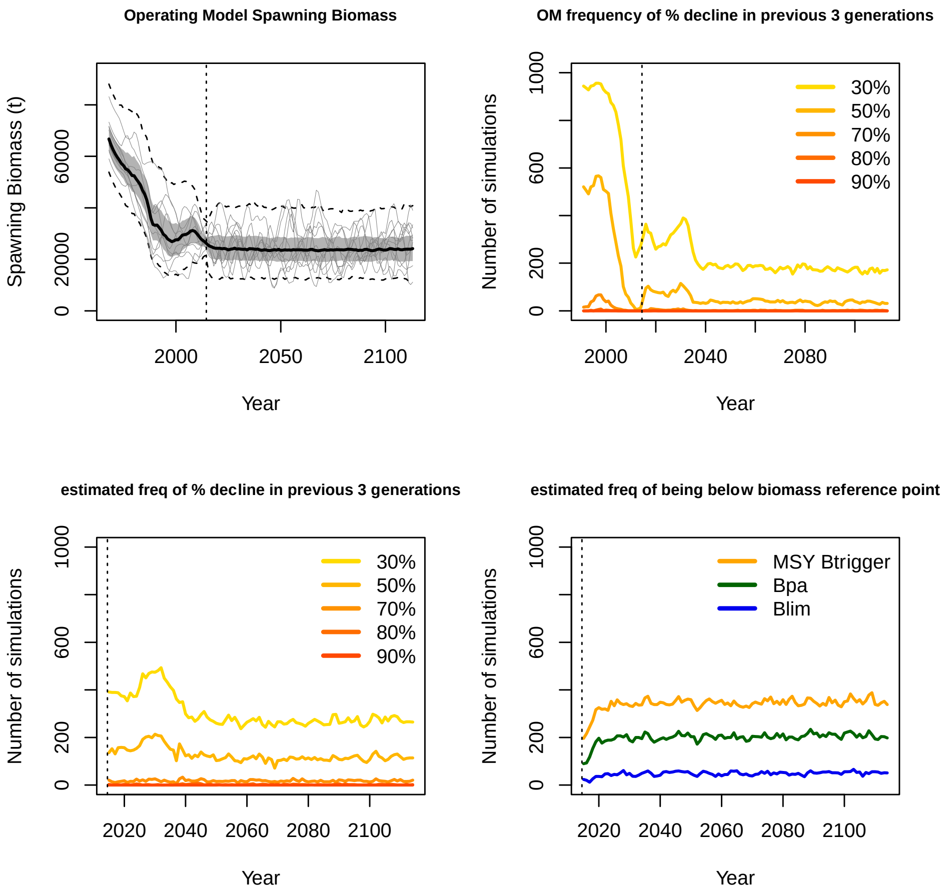


Figure 28: Time series of outputs for blue shark simulations when stock status at the beginning of the management projection period is 40% of unfished biomass. (a) summary of spawning biomass trajectory, with median (dark line), interquartile range (grey shading), and 95% interval (dashed lines). Example individual simulation trajectories shown in grey lines. (b) Number of simulations in which change in operating model spawning biomass over three previous generations exceeds thresholds for IUCN categories, (c) Number of simulations in which the estimated change in \spawning biomass over three previous generations exceeds thresholds for IUCN categories, (d) number of simulations in which spawning biomass is estimated to be below ICES biomass reference points.

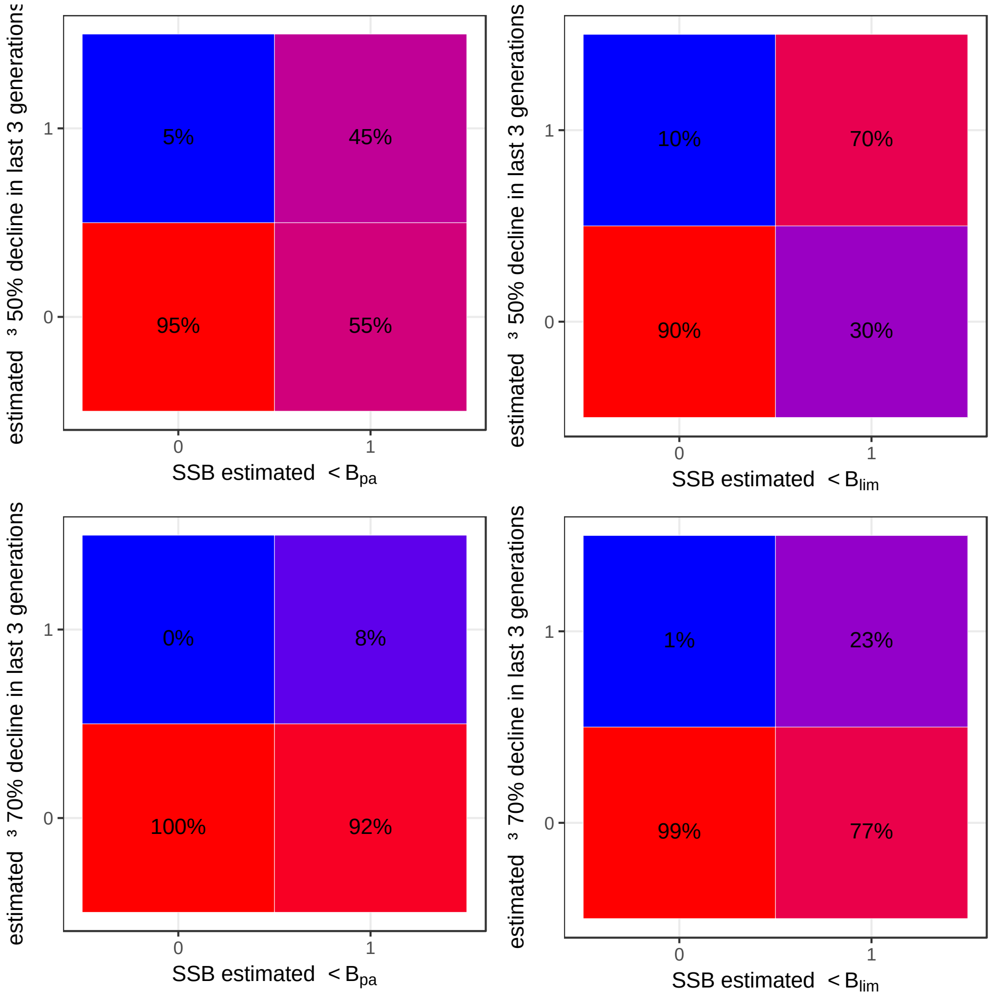


Figure 29: Confusion matrices for blue shark when stock status at the beginning of the management projection period is 40% of unfished biomass, with percentage of simulated assessments (over both projection years and simulations) where biomass change is estimated to trigger thresholds for IUCN criterion A (1) or not (0) given that spawning biomass is estimated to be below either B_{pa} or B_{lim} (1) or not (0).

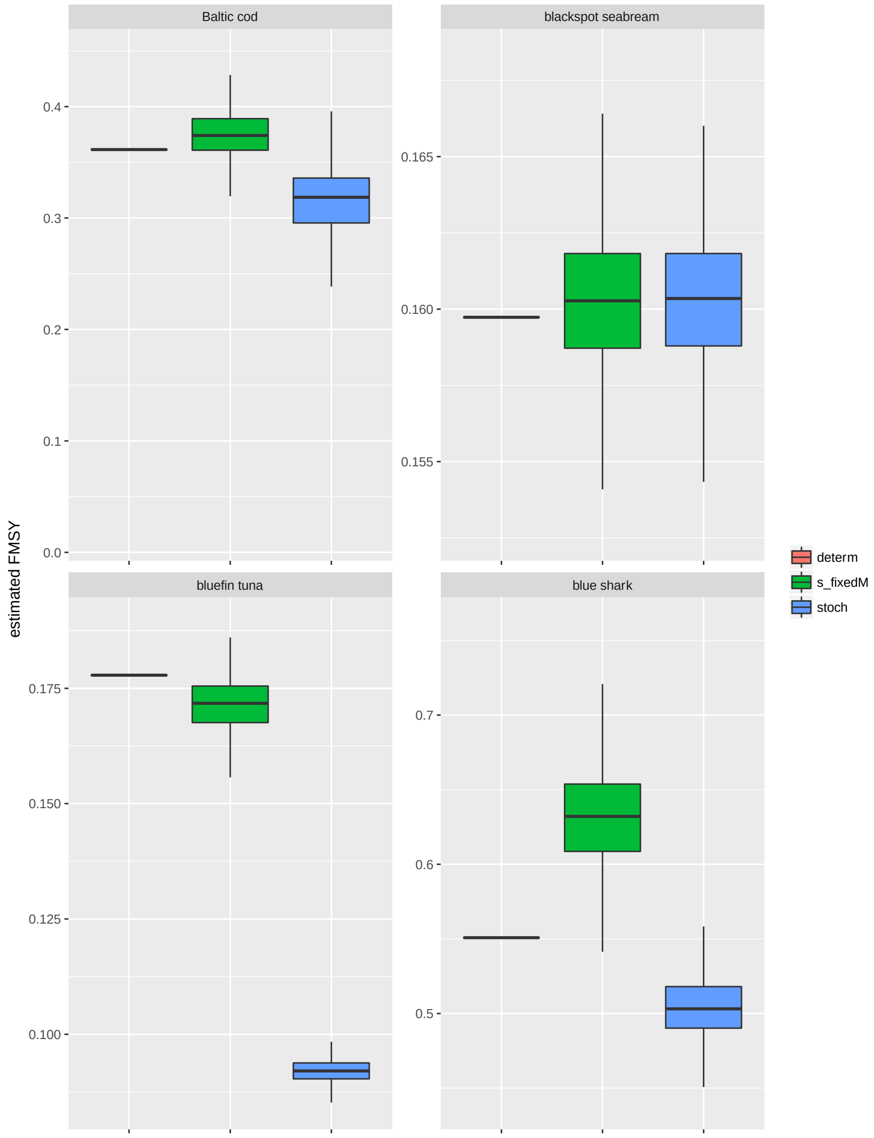


Figure 30: Estimates of FMSY for the four stocks obtained from: "determ" – the deterministic approach used within the MSE simulations, "stoch" – the stochastic projections from the OMs given assessment error and recruitment variability, attempting to mimic the ICES Advice method for defining reference points, and "s_fixedM" – as for the stochastic approach but the fixed M over age used in the deterministic approach is applied to the stock projections rather than the OM vector of M at age. Values for Baltic cod are the same for both western and eastern stocks.

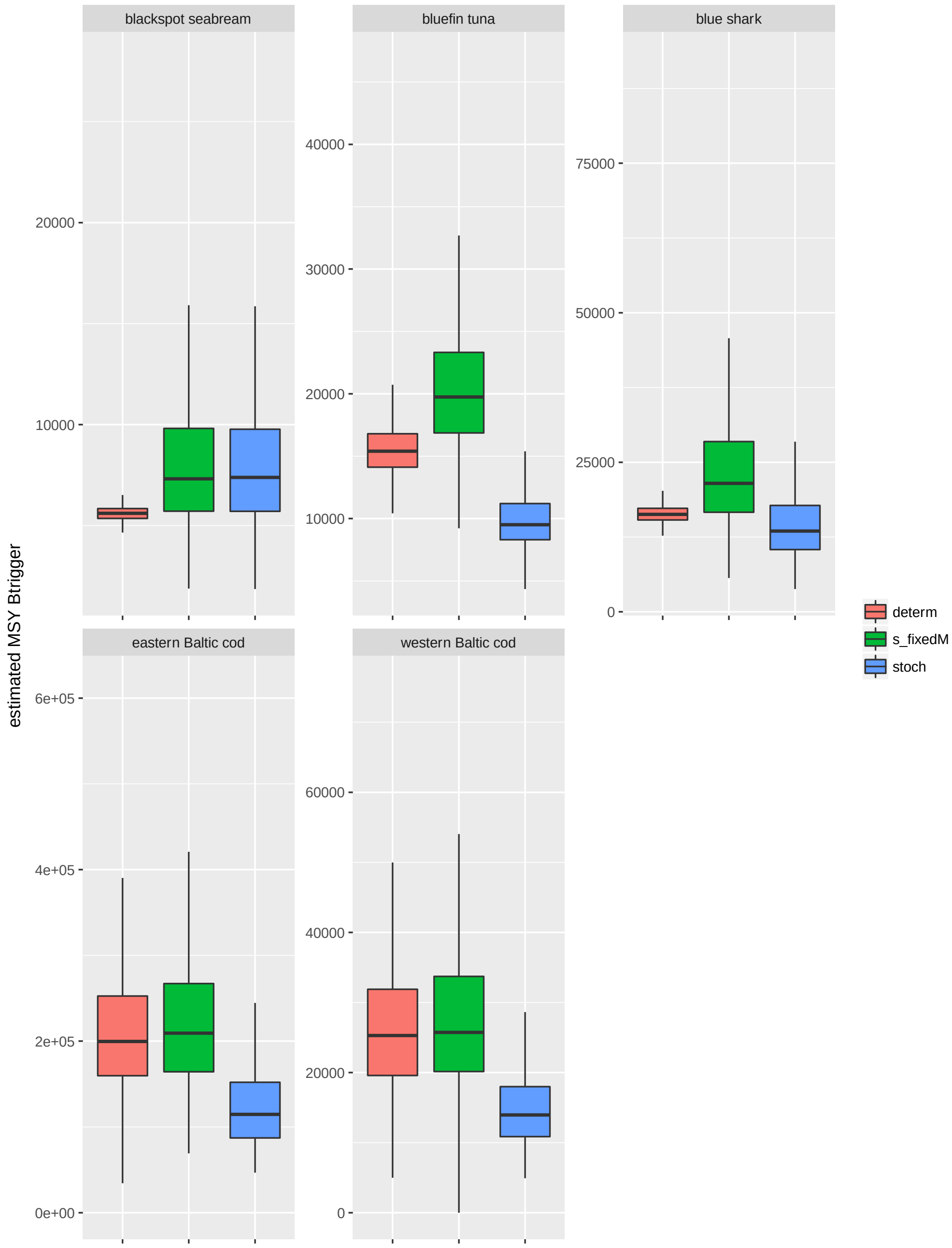


Figure 31: Estimates of MSY $B_{trigger}$ for the five stocks obtained from: "determ" – the deterministic approach used within the MSE simulations, "stoch" – the stochastic projections from the OMs given assessment error and recruitment variability, attempting to mimic the ICES Advice method for defining reference points, and "s_fixedM" – as for the stochastic approach but the fixed M over age used in the deterministic approach is applied to the stock projections rather than the OM vector of M at age.

Annex 3: Reviews

Report on IUCN assessments and fisheries management approaches Review

Joanne Morgan

There report presents two different approaches, which together attempt to provide the background to develop advice. I think that most of the aspects of the request are covered in the report.

One of the main conclusions is that the two approaches are complimentary. I am not sure that this conclusion is really valid. As the report points out, they are different approaches designed for different purposes. The fact that they sometimes 'agree' in their general conclusion about stock status does not make them complimentary. (For example compare Rice and Legace that found that 'the decline criterion suggested a serious risk-of-extinction in 87%, whereas most of the stocks were still within a zone that allowed fisheries management reference points to indicate that exploitation could continue' to the results of Fernandes *et al*. This comparison suggests that the approaches rather than being complimentary, they just sometimes agree and sometimes disagree). I think that the request can be adequately addressed without making this conclusion. The highlighted point should be 'offer differing insights and levels of information, and should not necessarily be taken as directly comparable'.

Below I go through each of the parts of the advice covered in section 9 and provide comment where necessary. Following this I also provide a few comments not directly related to specific conclusion sections.

- Perform a critical analysis of the IUCN assessment process for marine species in comparison the scientific process of relevant fisheries organisations (in particular ICES, scientific committees of RFMOs); including how data/information from different sources used by IUCN are weighted to derive a final perception of decline (section 9.1)

I did not find that this section addressed the weighting of different data/information sources used by IUCN directly. This may be because 'in practice commercial fish stocks are assessed only using criterion A' or because status is determined by the criterion that gives the highest threat category. This needs to be explicit in this section. In addition, if IUCN deals with cases where more than one set of data is relevant for a particular criterion, do they weight these data and if so how?

- Provide an analysis of the criteria used by IUCN and their suitability for marine fish species (section 9.2)

This section is quite clear. However, I would highlight here the fact that the IUCN criteria are less conservative than the methods used by ICES. This is covered in the next section but it seems to me that it is also relevant to the issue of suitability so it would be good to include it here.

- Provide a comparison between IUCN categories of conservation status and fisheries reference points for defining sustainability at both the species and stock level (section 9.3)

MSE is proposed as a 'more comprehensive tool for assessing uncertainty'. This is likely the case when a single population is being assessed for risk of extinction. However, it is less clear how this could be applied to a species.

- Identify pros and cons of the two processes

There is no specific section on this but it is covered generally covered in the report. A specific listing of pros and cons might help the ADG.

- Compare recent IUCN listings for key EU stocks with corresponding information from stock assessments. (section 9.4)

This item was very well covered and discussed in various places in the report. In this section I do question the use of the term 'general coincidence'. This seems to overstate things. Having 25% of stocks below B_{pa} but classed as non-threatened by IUCN approach seems an important 'non coincidence'. So I suggest using language like in section 7.3 '73% of stocks were similarly assigned'. I would also remove the reference to previous work (Davies and Baum) and in this section stick to the comparison done here. I also wonder about the comparison based on fishing mortality. Given that the IUCN approach is related to biomass trends it is not surprising that there would be more of a discrepancy when looking at fishing mortality. Is this really needed?

- Case studies species should cover a range of taxonomic levels and also include a range of data rich and data limited stocks. Case study species could usefully be agreed between ICES and MARE prior to any workshop.

Not specifically in section 9. See section 8 comments below.

- Assess suitability and reliability, and advantages of disadvantages of incorporating IUCN assessments for fisheries management and other relevant decision making processes. (section 9.5)

It is not really clear how 'coupling an IUCN assessment to a stock assessment' would work. Nor is it clear how this would 'supplement the knowledge base to assessing the CFP objectives'. MSE is suggested but it is not clear how this would work at the species level. There might need to be a change in focus of IUCN to populations instead of species to implement this.

The Norwegian approach is suggested as an example. This is not presented in enough detail to evaluate if it is a good idea or not.

Other sections

Section 3.2

The fact that 'whereas 13.5% of Red Listed marine fishes were classed as threatened, 40% of populations with a stock assessment were classed as being below biomass limit reference points' seems somewhat at odds with the conclusion that 'biological reference points for commercially exploited species and IUCN analysis were compatible and well aligned'

'In addition, the Powles *et al*. (2000) paper featured an incorrect and regularly-referenced figure that erroneously implied a fish stock/species would only get a threatened listing by IUCN when the stock had been completely collapsed by management, which is not the case.' There needs to be a reference for this or it needs to be removed. I didn't find that this statement really added anything.

'Rice and Legace (2007) undertook work similar to that of Dulvy *et al.*, but increased the number of species from 76 to 89, and used 35 years of population data. The authors found that assessing extinction risk using the IUCN criteria B (geographic range decline) and criteria C (number of mature individuals) were in harmony with limit reference points used by fisheries management agencies. Mangers should reduce fishing mortality to zero well before there was a chance of extinction. This study showed there was a large potential for conflict between fisheries and risk-of-extinction approaches when considering the extent of population declines (that is, criteria A).' This paragraph does not seem to adequately describe the results of the paper. It leaves me with the impression that they found things to be more or less in agreement whereas this was not the case.

Section 3.3

The timing of IUCN assessments relative to stock assessments may be important to highlight in the advice.

ICES is not a 'fisheries management' agency.

Section 5

The advice that was given on 'screening methods' is probably still relevant.

Section 7

7.2 I don't understand 'Reference points for these stocks were not always available but the ratios of B/BMSY and F/FMSY were, as well as the reference points.'

Section 8

8.2 'Populations were assumed to be unexploited in the initial years of the simulation', yet it seems that different OM started at different population levels relative to B0. Further the OM are not really well described in this summary. How are the OM started at different levels if conditioned on catch and survey?

The description of how performance is assessed is very unclear. This may partly be the result of the equations not being fully defined in the text. But there is also a lack of clarity in the text. For example 'In each year of the projection, the estimated SSB was compared to reference points and the magnitude of population decline over the previous 3 generations was calculated (as is required for calculation of IUCN criteria A).' Was the SSB from the true compared to the reference points from the true? What estimated SSB is and what reference points it was compared to (true or perceived) needs to be fully described. Similarly was criterion A calculated for the true or perceived? In all cases performance must be judged in the true world. If this has not been done then the results are not valid as presented.

One approach (again perhaps this was done but it is not clear) would be to look at performance in the perceived world and compare to performance in the true world. So look to see what threat status and what result for stock level vs reference points one gets using the perceived. Then judge if these are 'correct' or not by looking at what stock status and threat assignment that one gets in the true world.

8.3.1 'Only 1 out of 100 simulations…' should be 1000 I think.

8.3.2 'Effectively the larger (eastern) stock masked the smaller (western) stock, showing this method..' I think IUCN is being referred to here and this should be explicitly stated again in this sentence to be clear.

Peer Review

Report on IUCN assessments and fisheries management approaches

General

The report on IUCN assessments and fisheries management approaches addresses the request from the EU to ICES to undertake an "Analysis of the IUCN process for the assessment of the conservation status of marine species in comparison to the process used by fisheries management bodies" that should include the seven key objectives listed in the Introduction of the report.

The primary inputs and sources of information for the report were a comparison of recent classification of stocks (and species) falling within the EU according to the approaches and classification systems of the IUCN, ICES and ICCAT, which was undertaken by Dr P.G. Fernandes (Annex 1 of the report) and a simulation study that compared the results of assessments according to the IUCN and ICES criteria on simulated time series of four species, which was done by Dr G. Fay (Annex 2).

The two studies reported in Annexes 1 and 2 were well-designed and implemented and, in this reviewer's opinion, were sufficient and appropriate for evaluating and understanding the differences and different contributions of the conventional stock assessment approaches applied by ICES and ICCAT and the 'screening' method used by IUCN. The general conclusions from each of the two studies, and of the report as a whole (which I will not attempt to summarise here) are therefore valid and useful.

Nevertheless, in this reviewer's opinion, the report in its current form stops short of fully addressing the seventh key objective listed in the Introduction: to "Assess suitability and reliability, and advantages of disadvantages of incorporating IUCN assessments for fisheries management and other relevant decision making processes". There are frequent statements in the Annexes and the report itself that the two approaches (ICES and IUCN) are different, which is beyond dispute, and complementary, but without satisfactorily analysing how the IUCN approach could complement the ICES approach. In the context of this report, that potential complementarity is unclear and one of the primary messages coming out of the two studies is that "when a fish stock is exploited, the IUCN approach can underestimate the risk to a stock being outside safe biological limits" (section 9.3 and similar statements elsewhere). Further, the results reported in Annex 1 for biomass reference points show that the two methods yielded equivalent results for 61 stocks, the IUCN criteria generated what are referred to in the report as False Alarms for three stocks and for the remaining 20 stocks, the ICES approach assessed the stocks as being below the specified reference points (i.e. overfished) while the IUCN criteria did not detect a problem. The number of misses from the IUCN criteria increased to 33 when fishing mortality reference points were considered (i.e. whether or not overfishing is taking place). As a result, the report comments that 'the fisheries management reference points suggests that the IUCN approach is less precautionary than the fisheries management reference points'.

Further, from the results of the simulations in Annex 2, it was evident that "Towards the end of the 100 year stock projections, listing of Vulnerable again became more frequent, as populations were on average at higher levels, and moving around biomass targets". During those periods the stocks were being 'managed' through setting sustainable TACs, which, while the report does not explicitly state this, suggests that those Vulnerable listings were also equivalent to false alarms.

Based on these results and conclusions, it is not at all clear as to how the IUCN criteria can complement the ICES approach. However, there would appear to be other justifications for such complementarity that should be explored further, or at least emphasised more strongly in the conclusions of this report. These are:

i) Arguably, the most important complementary, and supplementary, role of the IUCN assessments is the much wider coverage of species by IUCN. In Annex 1 it is reported that for the most recent European Red List of Marine Fishes, IUCN assessments of extinction risk were carried out for 1022 species. In comparison, the ICES/ICCAT combined dataset used in the Annex 1 study covered 87 species. IUCN therefore greatly expands the number of marine species for which there is at least some information on status. It is likely that at least some of those additional species will be impacted by fisheries, directly or indirectly, and this information could be relevant to DG-MARE and other fishery authorities.

ii) The report refers to Norway as a country that directly addresses the IUCN categories and criteria in its fisheries management and it is reported that "The official Norwegian Red List for Species has become an important tool, or yardstick for management of economically important species. The aim is to minimise the risk of future listing of a species, and management measures are tailored for species already listed in order for that species to be removed from the list" (section 3.3). In the last paragraph of the main body of the report (section 9.5) it is stated that "The example of Norway, shows that the stock assessment and IUCN approach can be built together to provide a robust warning system for marine fish." It would be very useful to give more information in this report on Norway's approach to integrating the two approaches and thereby elaborate on the proposal that the two approaches should be seen as being complementary. In its present form, from the wording in section 3.3, it is unclear whether the IUCN listing brings additional, useful information to the management of fisheries or, as could be interpreted form the text, there is no actual integration but Norway simply aims to avoid to have its fisheries associated with undesirable IUCN categories and takes action to avoid this. This would mean, in effect, that improving the IUCN category is an objective for fisheries rather than a supplementary source of information and guidance. This could well be a misinterpretation and it is important that the report elaborates on just how Norway integrates the two approaches in its fisheries management.

iii) There is an unaddressed, and unavoidable limitation in the data coverage of this report, which seems attributable to effective management by the EU countries of their fisheries. As indicated by figures 7.1 and 7.2 (and their originals in Annex 1), there are only two species of the 84 assessed that were classified as Endangered and three as Vulnerable when comparing against biomass (fig 7.1) or two Critically Endangered, one Endangered and 4 Vulnerable when comparing against fishing mortality (fig 7.2). The rest are classified as being at lower risk (Near Threatened or better). This means that most of the comparisons done here, in which the stock assessments were found to be frequently more precautionary, were against species of limited conservation concern. The question as to which approach would be more effective and precautionary in detecting the higher levels of conservation concern remains open at present. In principle, one would still expect it to be the more intensive stock assessment approaches but perhaps the existing set reference points would need to be extended to be able to classify those higher levels of risk. Without that, the IUCN classification should be important in identifying the species of greatest concern.

Some more specific comments

Glossary. The report assumes that readers will be familiar with all the abbreviations and terms used in this report. I cannot comment on whether that will be the case or not, but the authors may want to consider whether the addition of a glossary would make the report easier to read and understand for a wider audience.

Reference points

Clear explanations of what B_{lim} and B_{pa} are and how they are calculated need to be provided, as well as the rationale for developing these reference points in the first place. This will help the readers to compare those reference points, and their intentions, with the IUCN criteria. They are defined in section 2.5 of Appendix 2 (B_{lim} = 0.3 B_{MSY} , and B_{pa} = 1.4 x B_{lim}) but should also be in the main report and the rationale for using those points explained.

Sections 6.3 and 6.7 – In order to strengthen compatibility, would it be feasible and would there be value in trying to develop reference points for stock assessment approaches that would equate to the IUCN vulnerability criteria or would the fundamentally different approaches make this impractical or simply not feasible?

Section 6.4 – 'and habitat relationships'. This reviewer's experience is that taking the temporal variability of the environment into account is still the exception rather than the rule in stock assessments generally. However, there are reasons for the failure to use, or infrequent use of, habitat and environmental change or variability in stock assessments, which are related to typically high levels of uncertainty and limited confidence in most available estimates of relationships between these variables and population dynamics variables. A question on the IUCN criteria is therefore whether and how these uncertainties are dealt with there.

Section 8. Check the figure numbers in this section as some of them were incorrect in the version I was sent.

Section 8.3.3. Under this section the example of two stocks of Baltic cod is presented, for which the results showed that when the IUCN biomass decline statistics were calculated from the eastern and western assessment estimates combined, cases in which the >=50% threshold was triggered was low and that the biomass of the larger eastern stock dominated that of the smaller western. The report concludes that an approach of combining fisheries stocks for assessment against the IUCN criteria may not be appropriate when a smaller stock is in worse condition that a larger stock. I suggest that this result could be generalised to cases where there are more than two stocks, particularly in cases where some of the stocks make only relatively small contributions to total species biomass. By extrapolation of this principle, the IUCN species-wide approach would therefore not be appropriate in cases where there is a large enough number of stocks that all or most of them make only small contributions to total biomass.

Section 8.3.4. This section presents the results for blackspot seabream from Appendix 2. In its current form, no reference is made to the comparisons between the results of the IUCN and ICES approaches (Appendix 2, Figure 20 and associated text). The data poor status of this stock and the resulting high CV in the pseudo assessments have an impact, albeit it small, on the match-mismatch performance of the two approaches (see Figure 20), which should be reported in this section.

Section 8.5 (also 9.3). In this section (2nd paragraph) it is noted that "Towards the end of the 100 year stock projections, listing of Vulnerable again became more frequent, as populations were on average at higher levels, and moving around biomass targets." This is an important result that should be discussed further. On the basis of this statement, and the results that prompt it, it would seem that false alarms could be a problem in an adequately managed stock but that undergoes sufficient variability to lead to sustained declines that would be detected by the IUCN criteria. The validity and implications of this observation should be discussed.

Section 9.3. In the 2nd paragraph of this section it is stated that "Simulations suggest that the IUCN threatened status will be infrequently triggered in a stock that is stable below biomass reference points, due to a lack of further decline in population size. ……. This is a result of the differing objectives of each system." I disagree with that explanation and in my view this result demonstrates a weakness in the IUCN criteria. A population or stock that has been reduced to a low level, such as below B_{pa} is likely to be at higher risk of extinction than one that is well above such a limit reference point and one would expect the IUCN criteria to be able to reflect that. As an example, the male biomass of the South African West Coast rock lobster population *Jasus lalandii* is currently at less than two percent of its unexploited biomass. Efforts are made to harvest it sustainably at this very depleted level, for social reasons, and it the population is (or should be) kept more or less stable, which would mean it would likely to be given a low risk status by the current IUCN criteria. Nevertheless, the risk of further decline and extinction must be substantially higher than if the biomass of the population was considerably higher and I would argue that this heightened risk should be reflected in an IUCN (or any other) classification.

Section 9.3, 4th paragraph. It is stated here that 'Simulated stocks very rarely, or never, fell below the 10% spawning biomass threshold.,,,,,, Whereas an assessment using criterion A (population decline) would have led to a listing for a few simulated populations. Taking into account the warning about evidentiary attitude in the IUCN guidelines, these few simulated populations would be listed, although there was no risk to the populations becoming extinct, under assumed conditions." This argument ignores the fact that the IUCN categories are intended to highlight different degrees of threat or risk. With such an approach, there could be justification for giving early alerts if the risk of falling below thresholds higher than 10%, which is also why ICES and the CFP use thresholds such as B_{pa}.

Section 9.4. It is stated that "When using fishing mortality reference points, it appears that the IUCN assessment has a reduced ability to alert to over exploitation. This highlights that as a conservation tool, the IUCN approach is not designed to advise on maximising yield." Again, I disagree with this conclusion (as did Punt (2000) who you refer to in the next sentence). These misses by IUCN have nothing to do with sustainable yield and (as also stated in my comment on section 9.3 above) the fact that the IUCN criteria do not pick up relatively stable populations but at levels below limit reference points demonstrates a weakness in the IUCN criteria from the perspective of monitoring conservation status, because they will not detect such populations even though their conservation status is likely to be less secure than if they were at higher biomasses.

Section 9.5. In the 3rd paragraph it is proposed "For commercially exploited fish, this can be easily achieved by coupling an IUCN assessment to a stock assessment. This would supplement the knowledge base to assessing the CFP objectives of long term sustainability and minimising negative impacts of fishing activities." This is probably a good way forward but, as argued in my General comments, the authors need to explain in some detail what they have in mind.

Appendix 1.

Methods. It is not clear (unless I missed it), which years were used for the ICES and ICCAT assessments in comparing with the 2015 IUCN assessments. In the Discussion, when examining possible reasons for false alarms, the author suggests one explanation could be a mismatch in the timing of either assessment. It needs to be stated under Methods whether the ICES or ICCAT assessment done as close to the latest IUCN assessment as was available were used for the comparison (which would have been the best option), or the most recent assessment or some other selection.

Detailed comments

Additional comments and some suggested editorial corrections are shown in the copy of the draft report that is also attached to the covering email. Some of these overlap with the 'more specific comments' provided in the previous section.

Kevern Cochrane
Department of Ichthyology and Fisheries Science
Rhodes University
South Africa

18 October 2018.