ICES Technical Service
EU ecoregions
Published 12 November 2021

EU request for a Technical Service on MSFD Article 8 guidance on undertaking assessments for Descriptor 3 (commercially exploited fish and shellfish) and Descriptor 4 (marine foodwebs)

Service summary

ICES provides guidance for assessing Descriptor 3 (D3; healthy populations of commercially exploited fish and shellfish) and Descriptor 4 (D4; abundance and reproduction of marine foodweb elements) under the EU Marine Strategy Framework Directive (MSFD) by preparing the Article 8 guidance of Commission Decision for D3 and D4 2017/848 (EU, 2017).

The guidance document has been prepared by ICES appointed experts, selected both to cover Europe geographically and for their experience in the assessment of commercial stocks and foodwebs, thus facilitating the identification of commonalities and synergies for the guidance.

Existing documents on the topic of D3 and D4 were reviewed and analysed (i.e. a JRC review of Member State reporting [Vasilakopoulos *et al*., 2021; Boschetti *et al*., 2021], ICES workshop reports, and various RSC reports including the recent HELCOM ComFish 1-2021 [HELCOM, 2021]). This technical service summarizes the main state of play, priority actions, and suggestions for both D3 and D4. Article 8 guidance, see annexes 1 and 2, is prepared using templates agreed-upon within the Common Implementation Strategy process of the MSFD.

For D3, the guidance provides further specifications for existing approaches for assessing criteria 1 and 2 (C1 and C2). These specifically relate to the need for regional coordination, the evaluation of status in the six-year MSFD assessment period, the assessment of stocks for which data and/or threshold estimates are not currently available, and the integration of criteria assessment results to attain stock status. For Criterion 3 (C3), the guidance recommends an expansion of the current interpretation to adequately monitor the development in stock productive capacity and ensure alignment with assessments conducted under Descriptor 1 (maintaining biodiversity). The guidance recommends that indicators under Criterion 3 can be assessed with both lower and upper thresholds and that classification on Criterion 3 below Good Environmental Status (GES) trigger a re-evaluation of thresholds under criteria 1 and 2. However, in some cases, it may also be appropriate to consider a non-threshold approach. Further work is required to derive detailed threshold guidelines for Criterion 3 indicators and to ensure that these thresholds are consistent with those under criteria 1 and 2 and the overall objectives of D3.

For D4, the guidance suggests common trophic guild definitions in order to resolve the current difficulties in comparing assessments between EU Member States and across regions. Temporal consistency in assessments both within and between descriptors is particularly important for D4; such guidance can most appropriately be agreed upon at Commission level. Once an appropriate spatial scale of the foodweb assessment has been identified based on the spatial distribution of its core elements, the foodweb is assigned to one of the features 'coastal', 'shelf', 'slope', or 'deep' waters. Details on how to assess each of the four criteria under D4 are given with the aim of providing greater coherence and comparability between the assessments in different areas and suggesting indicators that have demonstrated consistent performance. The development of methods to integrate assessment results across the different guilds to assess the balance of the foodweb is still not at a stage where tested and peer-reviewed indicators are commonly agreed upon. This is a key issue that needs to be addressed before assessing the status of the foodweb from the guilds monitored. The lack of agreed integration methods should not lead Member States to refrain from reporting the status of individual guilds.

Indicators of criteria under D3 and D4 can be assessed using a combination of models or simpler data-based indicators. In all cases, methods should be quality controlled and peer-reviewed. There is a need for cross-descriptor guidance, agreed upon at the EC, on which years to include to ensure comparability between descriptors (and regions). The impact of climate change in the indicator-specific setting of thresholds for future iterations of D3 and D4 also needs to be considered.

Request

EU (DG ENV) requests ICES to provide a Technical Service on MSFD Article 8 Guidance on undertaking assessments for: i) Descriptor 3: Commercially exploited fish and shellfish; and ii) Descriptor 4: Marine food webs. The guidance should

compile available information on how to undertake an assessment, and identify where there are gaps or ongoing work (e.g. development of lists for D3). For D3, the methodology for assessments is quite mature, being based on CFP assessment procedures for D3C1 and D3C2, and drawing from recent ICES advice on lists of species and D3C3 indicators. For D4, there is a need to review latest approaches by member states (2018 MSFD reports) and regional sea conventions (ongoing work for 2023 quality status reports) and develop a common EU framework.

ICES to make an early draft of D3 and D4 Article 8 guidance available to DGENV so that DGENV can provide feedback to ICES on:
- DG GES input as overall coordinator for the drafting, in particular the general chapters in the guidance document
- Milieu input as administrative coordinator, in particular for hosting the documents on their Microsoft Teams platform, compiling contributions and ensuring guidance homogeneity across chapters
- DG ENV input as responsible for the timely and quality delivery of the guidance document, to be adopted by WG GES

Main state of play

There was substantial variation in Member State reporting for the latest assessment of D3 and D4. There were differences within and between regions in the interpretation of the guidelines—specifically on the selection of stocks under D3 and guilds under D4, the time period included in the assessment, the foodweb categorization of features, and the indicators and thresholds used. Furthermore, few criteria were consistently monitored under D4, and very few of them had associated thresholds. It was also unclear in the guidance how consistency between linked descriptors should be attained. For D4, the current definition of trophic guilds lacks detail, and the current flexibility in guild definition makes it difficult to compare assessments across regions.

Priority actions

For both D3 and D4, there is a need for cross-descriptor guidance on which years to include to ensure comparability between descriptors (and regions). Guidance can most appropriately be agreed upon at Commission level. Similarly, the spatial scale most appropriate for foodwebs in a specific (sub)region should be agreed upon regionally.

Suggestions

Further work is needed to agree upon at EU level the years to include in descriptor assessments and at a regional level the spatial scale of the foodweb(s) assessed, as well as appropriate guilds to monitor and assess within these assessments. Agreements should ideally be performed regionally and as a minimum between Member State sharing the same foodweb assessment area level on indicators and thresholds appropriate for monitoring D3 Criterion 3 and all D4 criteria. The years should be consistent between descriptors D1, D3 and D4 as a minimum. To support this, it is suggested that dedicated international workshops are conducted to advance the identification of indicators and thresholds for these D3 and D4 criteria; this work should include the integration of approaches to attain measures of balance between guilds under D4. A lack of agreed-upon integrated approaches should not result in a lack of reporting on criteria indicators.

Several examples exist of interactions between assessment elements of D4 and descriptors 1, 2, and 3 that are of relevance for management towards MSFD objectives. Observed ecosystem and foodweb changes potentially provide information that can be used to improve the predictions of future stock productivity under D3. For example, an increased population of seals is likely to result in a sustained elevated natural mortality of its prey. Changes to the abundance and composition of zooplankton, as often anticipated under climate change, is likely to impact the productivity of planktivorous fish, including the larval stages of most of the commercially exploited stocks. Species that are introduced or are migrating into a new area may increase in abundance to the level where they significantly alter the foodweb, either increasing predation, decreasing food abundance, or providing a new alternative food source.

Background and approach for the request

The EC request was for ICES to draft an Article 8 guidance document for the assessment of MSFD D3 and D4. The guidance document has been prepared by ICES appointed experts, selected to cover Europe geographically from north to south and for their experience in the assessment of both commercial stocks and foodwebs. The task was approached

through compiling and reviewing existing documents on the topic of D3 and D4, including the recent report from HELCOM (HELCOM, 2021) which was published while the guidance was being worked on. Completing the guidance documents for the two descriptors in parallel facilitated the identification of commonalities and synergies in the assessment guidance for both D3 and D4.

The comparison or transfer of results between different indicators relies on the underlying indicators reflecting comparable spatiotemporal scales. If this requirement is not fulfilled, the foodweb can *appear* to be out of balance due to spatiotemporal differences rather than show actual imbalance. In the most recent assessment, there was little consistency between the temporal scales of different assessments. The years included in the 2018 assessment of D4 varied between 1958–2014 and 2013–2018. There were pronounced regional differences as the Baltic assessments mostly used 2011–2016, the Northeast Atlantic mostly used 2012–2017, and the Mediterranean mostly used 2013–2018. There needs to be cross-descriptor guidance on years to include to ensure comparability between descriptors (and regions). Such guidance can most appropriately be agreed at Commission level and should balance the need for up-to-date information with the time taken to estimate and quality control indicator values and predictions. The spatial scale of the assessments performed under D1, D2, D3, and D5 differs both within and between descriptors, making it necessary for D4 to accommodate differences in assessment area.

The MSFD requires EU Member States to account for climate or environmental change in the indicators and/or the setting of thresholds. It is noted that with the large recent changes in climate and ecosystems, this is a key issue to address in future studies. Under both D3 and D4, it was decided not to make recommendations on methods to integrate across stocks/guilds, as this was considered beyond the remit of the technical service request.

Background of D3 guidance

Approach

The JRC review of Member State reporting on D3 was analysed (Vasilakopoulos *et al*., 2021), and it was noted that there are often substantial differences in approaches used in the ICES and GFCM areas. It was therefore made a priority to ensure that the guidance would be useful throughout European seas. Furthermore, defining indicators for GES that were consistent between the MSFD, CFP, and multiannual plans and that allowed the monitoring of stock quality and productivity while retaining comparability with Criterion 3 of Descriptor 1 (D1C3; HELCOM, 2021) was considered a key aspect.

Stocks included

In 2020, ICES provided advice on how to select stocks for inclusion in D3 reporting (ICES, 2020). This advice will be transformed into operational guidance once there is common acceptance of the best approach to prepare the lists for D3. Option 1 in the advice states that all Member States will report on the same species for a region using EU level assessments, though individual Member States can choose to add stocks of national relevance. There is limited guidance on how to choose stocks of national relevance, and there is no general agreement on the proportion of the catches that must be covered by the selected species in a region. Further, it is not detailed how stocks straddling several regions will be included. On top of that, species that have been important landed species in previous MSFD reporting periods but are currently at very low abundance should be retained on the list.

A stock should not be assessed both under D1 and D3 using different criteria. Hence, if the stock is included in D3, there will be no new assessment under D1. However, the status of the stock under the D3 criteria (D3C1, D3C2, and D3C3) can be used in species-based (D3 assessment integrated across criteria and exported to D1) or criteria-based integrations (criteria-wise import of D3C1 to D1C1, D3C2 to D1C2, and D3C3 to D1C3) under D1. It was noted that there is a potential for different stocks of a species to emerge in both D1 and D3; for example, if a sensitive species in an area is divided into several stocks of which only some provide large contributions to landings (e.g. anglerfishes and rays).

Assessment of D3C1 and D3C2

Considerations for D3C1 and D3C2 for stocks with assessment and MSY reference points included the issue of ensuring consistency between the CFP, agreed multiannual plans (MAPs), and the implementation of D3C1 and D3C2. The CFP

and the MAPs both refer to F_{MSY} as a target and to biomass safeguards as biomass levels below which stock productivity is impaired and measures should be taken to rebuild the stock to levels of full productivity. The use of F_{MSY} as a target means that in an unbiased management system where the prediction and implementation of F is almost perfect, the threshold F_{MSY} will still be slightly exceeded in half of the years.

As a further complication, the MAPs also define F_{MSY} ranges of which the upper ranges can be used if the stock is above the agreed precautionary reference level B_{pa} and if the following criteria are met:

- on the basis of scientific advice or evidence, it is required that the objectives laid down in Article 3 in the case of mixed fisheries are achieved;
- on the basis of scientific advice or evidence, it is required to avoid serious harm to a stock caused by intra- or inter-species stock dynamics; or
- in order to limit variations in fishing opportunities between consecutive years to not more than 20%.

Fishing mortalities above F_{MSY} are hence in certain cases in accordance with the MAPs. However, the MAPs also state that the plan aims to 'ensure that the conditions described in Descriptor 3 contained in Annex I to Directive 2008/56/EC[1] are fulfilled (and other available MAPs[2]). While neither of the two texts is clear on the temporal scale for assessing this, it would follow logically that F should not exceed F_{MSY} on average in the MSFD assessment period though this may occur in individual years.

The MAPs further define limits to biomass (biomass safeguards) below which measures should be taken to reduce fishing impact. These biomass safeguards are often used as thresholds for stock biomass. However, for some stocks (e.g. short-lived species), these safeguards are target levels which ensure that there is a 95% probability of not adversely affecting recruitment. For other stocks, there is only a requirement to be above the biomass threshold on average and therefore no guarantee that the biomass thresholds will be avoided in more than 50% of the years. These considerations would be true even if there were an extremely good understanding of stock dynamics and extremely efficient management. Again, even for these stocks, the objective of fulfilling the Descriptor 3 requirements would indicate that as a minimum, stocks should be above biomass reference levels on average in the MSFD assessment period even if the stock may not meet the requirements in an individual year.

The current guidance for D3 does not explicitly address the incorporation of productivity changes caused by ecosystem changes. However, a stock with decreasing productivity due to low growth or increased mortality cannot necessarily sustain the same fishing pressure as a more productive stock, and the thresholds defined for such a stock should reflect this. In some cases, predation from species monitored under D1 may pose additional mortality on the stock and this link between D1 and D4. In both cases, an update of thresholds may be needed.

A further issue is the lack of agreement on which threshold values should be used for the assessment. For example, F_{MSY} varies over time and is occasionally updated to match current productivity conditions in line with the guidance for the MSFD. This new F_{MSY} pertains to current conditions but not historical conditions, and using the value for the historical time period provides a biased view of the development of fishing pressure relative to F_{MSY}. Stocks where F_{MSY} has increased will appear historically underexploited and vice versa. Other updates are also frequently made, meaning it is difficult to derive historically comparable thresholds for F_{MSY} from the literature. Ideally, this could be addressed by providing updated estimates of F_{MSY} for all years when updating assessment methods.

Several of the exploited stocks in the EU remain below levels at which recruitment is considered impaired (B_{lim}) despite of stated policy aiming to remedy this. For such stocks, the CFP, MAPs, and scientific advice all agree that measures should be taken to reduce F. While the method to attain this goal is not specified, the F threshold for such stocks must be lower than F_{MSY} to fulfil the requirement for reduced pressure below MSY $B_{trigger}$. For stocks where F is regulated through effort controls rather than through TACs, it is key that these effort controls are efficient.

Species which do not have full analytical assessments and associated thresholds are generally depicted as data deficient. These data deficient stocks include those impacted lightly as well as those heavily impacted but now at too

[1] https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R1139&from=en.
[2] https://ec.europa.eu/oceans-and-fisheries/fisheries/rules/multiannual-plans_en.

low levels to allow sufficient data to be collected. This opens the opportunity for overfished stocks to be excluded from D3 assessments due to data shortage as they become scarce.

Species which do not have full analytical assessments and associated reference points but occur in the stock list derived from the approach of ICES advice on D3 lists (ICES, 2020) may not have historically been considered a priority for assessment. However, with their appearance on the D3, data should be assembled at a regional level and attempts made to construct an assessment model which can be used to assess the status of the stock relative to MSY reference levels.

Species which are severely depleted compared to historical reference levels could be ranked as 'biomass depleted below last estimated reference level' with an associated estimated confidence in that rating depending on the data available. For example, if survey catches of the species have not increased, it is unlikely that the stock has rebuilt. If the species is still caught in sufficient numbers and has a reliable biomass index from surveys, an index of fishing mortality can be derived as catch divided by the biomass index for the period before and after the latest survey.

Possible reasons for not meeting thresholds were also discussed. There are two major reasons why thresholds are not met on average over several years: the governance system does not implement and control sufficiently restrictive fishing opportunities and the scientific advisory system provides inaccurate advice.

The governance system may fail either because of a lack of international agreement on shared stocks (EU and non-EU total catch allowance exceeds the scientific advice) or because the EU total allowable catch is inefficiently implemented (e.g. through a non-enforced limit to fishing effort or a lack of efficient implementation of the landing obligation). In these cases, giving the stock a red status can help create public awareness and pressure for greater adherence to scientific advice.

The scientific advisory system may provide inaccurate advice if recent changes to stock productivity lead to biased stock assessments and these again lead to too high or too low catch advice. This can happen, for example, if the stock experiences greater than expected mortality or lower than expected growth and recruitment, situations observed in several stocks in recent years. If these situations persist, they should result in an update of the thresholds for D3C1 and D3C2. Such updates may lead to a downscaling of F_{MSY}, and this downscaling automatically results in the F threshold having been exceeded in previous years where a higher target F was used in advice. In these cases, governance actions may have been appropriate and sufficient according to the scientific advice—but still the stock was overexploited. It would be informative to include some information on the likely role of each of these factors along with the confidence in the assessment.

Assessment of D3C3

A variety of indicators can be used to assess D3C3 although no specific methodological approaches and thresholds are currently agreed upon. The criteria used under D1 and D3 are very similar for criteria 1 and 2 but diverge for criterion 3, where D3 is considerably narrower than D1C3. An opportunity for broadening these, however, is given: 'D3C3 shall reflect that healthy populations of species are characterized by a high proportion of old, large individuals. The relevant properties are the following: (i) size distribution of individuals in the population, expressed as: — the proportion of fish larger than mean size of first sexual maturation, or — the 95th percentile of the fish-length distribution of each population, as observed in research vessel or other surveys; (ii) genetic effects of exploitation of the species, such as size at first sexual maturation, where appropriate and feasible. Other expressions of the relevant properties may be used following further scientific and technical development of this criterion.' (EU, 2017). Compared to: 'D1C3 —The population demographic characteristics (e.g. body size or age class structure, sex ratio, fecundity, and survival rates) of the species are indicative of a healthy population which is not adversely affected due to anthropogenic pressures. Member States shall establish threshold values for specified characteristics of each species through regional or subregional cooperation, taking account of adverse effects on their health derived from D8C2, D8C4 and other relevant pressures.' (EU, 2017).

Given the observed changes in stock productivity, ICES recommends to expand the focus of D3C3 beyond age composition and genetic diversity and to account for other properties. A list of potential indicators using an expanded approach for the assessment of D3C3 is given in Annex 1.

Background of D4 guidance

Approach

ICES analysed the JRC review of Member State reporting on D4 (Boschetti *et al.*, 2021) and noted that there are often substantial differences in approaches used in the ICES and GFCM areas. It was therefore a priority to ensure that the guidance would be useful throughout European seas. Further, there is a need to define consistent guilds and reconcile spatial differences in assessments, both considered key aspects.

Spatial coverage of foodweb indicators under D4

Reconciling differences in spatial scale between ecologically relevant assessment and reporting areas is also a cross-descriptor issue. This issue is particularly important when integrating components in assessments of foodwebs. This is because a foodweb needs to be relatively well-defined before the monitoring of it provides results that reflect the foodweb more than the natural variation between years in migration timing and extent. Marine foodwebs often have diffuse boundaries and substantial (often seasonal) migration over the borders of any area division.

In conclusion, the spatial scale at which foodwebs are assessed will only be the scale relevant to all stocks in exceptional cases. Instead, a spatial scale needs to be selected that reflects the foodweb in a specific subregion and is relevant to at least some other components monitored in this subregion. Ideally, the foodweb assessment area chosen should either have limited migration or good information on the timing and extent of migration.

Foodweb elements generally ignore EEZ boundaries except in cases where these elements coincide with boundaries such as very deep water. Combined with the spatial requirements, this means that only in rare cases is it appropriate to assess foodwebs in a Member State EEZ. It therefore follows that foodwebs should be assessed at regional level in regionally coordinated actions unless there are scientific investigations showing that this scale is not relevant to the trophic guilds and their links to other descriptors. In cases where subregions are spatially segregated from other subregions, the assessment can be done at a subregional scale. If a Member State collects information to suggest that smaller areas are consistent with these requirements, additional assessments can be provided for smaller foodwebs (e.g. coastal foodwebs).

Mapping foodwebs and guilds to ecosystem features

Having first agreed on the spatial scale of the foodweb, assigning this foodweb and the guilds within it to depth features is straightforward as long as these features do not overlap. To ensure this, each feature should be changed to one of 'coastal' (mean depth 0–50 m), 'shelf' (50–200 m), 'slope' (200–800 m), or 'deep' (> 800 m) waters. The foodweb assessment should be assigned based on mean depth in the spatial area covered using these categories.

Consistent trophic guilds

The current definition of trophic guilds lacks detail, and the current flexibility in guild definition makes it difficult to compare assessments across regions. To remedy this, ICES defined guilds expected to be relevant in all areas. The development of methods to integrate across the different guilds to assess the balance of the foodweb is still not at a stage where tested and peer-reviewed indicators are commonly agreed upon. This is a key issue that needs to be addressed before assessing the status of the foodweb from the guilds monitored. It should be noted that the evaluation of the balance is highly dependent on the guilds selected for monitoring: if guilds are several trophic levels apart, natural variability in the flow between levels may make balance measures difficult to interpret (e.g. estimating the balance between primary producers and sub-apex predators). The lack of agreed-upon integration methods should not be used as an excuse to refrain from reporting the status of individual guilds.

Indicators for the assessment of the criteria and threshold values

ICES noted that there is often a perception that foodweb indicators either should not or must not be model-based. This was considered a misperception – the issue is not whether a criterion is assessed using models or simpler, data-based indicators but rather that any method used to assess a criterion is thoroughly peer-reviewed and quality controlled.

For example, models are integral in the assessment for D3 as is their quality control through regular peer-reviewed benchmark processes. There is no reason that the use of models should not be allowed for D4 as long as similar standards are followed.

All indicators are assessed separately for each guild. As the indicators can be highly affected by natural variation, the thresholds should be defined as upper and lower levels reflecting natural variability and the ability of the indicator to reflect structural changes over time.

Indicators for D4C1

Indicators under this criterion reflect the relative abundance of species (also known as species dominance) or the relative change in abundance of species. As the former indicators reflect changes in the balance between species rather than just changes in their abundance, they are considered preferable to the latter where sufficient data are available. Examples include diversity indices such as the Simpson or Shannon indices and indicators of the proportion of seabirds that are increasing (OSPAR/HELCOM/ICES JWGBIRD, 2020). Where species information is not available, the indicator D4C3 using information on size should be monitored instead. Where only species groups are available, data at the lowest available taxonomic level should be used.

When referring to abundance, this should preferably be measured as biomass to ensure consistency with D4C2. Where several stocks or assessed populations of a species reside within the foodweb assessment area, these stocks or populations are treated as separate entities in the diversity analysis.

Indicators for D4C2

The derivation of estimated indicators for biomass is described in detail in the updated guidance.

Indicators for D4C3

Size-based indicators are reliant on information on the size distribution of the community, and the estimated indicator is highly dependent on the size selectivity of the method used to derive the data. The size distribution of a guild reflects species composition, growth, and total mortality. Therefore, the indicators used here depend on the indicators under the other criteria. Potential data include information from commercial catches, commercial landings, survey catch rates, and age- or size-based stock assessments (age-based assessments must be combined with information on length distribution-at-age). Among these, survey catch rates and stock assessments are considered most reliable, as commercial catches are based on a variety of fishing gears, each with its own size selectivity. Commercial landings are not considered appropriate due to the unknown quantities of discards of particularly smaller fish. In cases where no information is available on the size distribution, the mean maximum length indicator (MLI) can be applied.

Examples applied for fish and zooplankton include the large fish indicator (LFI), typical/median/mean/95th percentile of length (including the HELCOM core indicator zooplankton mean size and core stock, which is partially applicable as it integrates distribution and biomass), and mean maximum length (MML). Among this list of indicators, median or 95th percentile length is preferred over mean length because of the greater stability of the median when applied to highly skewed data. When using the LFI, the size of a 'large fish' needs to be estimated separately by guild and region, and the indicator is therefore not directly comparable between regions.

Mean maximum length (MML) integrates aspects of species diversity and size structure (mean possible length in the guild) and has the advantage that it can be estimated without information on size distribution of individuals in the guild. It is estimated using mean asymptotic length or a proxy for this for the species and species biomass. However, it is less responsive to changes in mean size in the guild than median length. If used, the reference for the maximum length used in the study should be clearly given, and the maximum length should be constant over time within a region.

When referring to abundance, this should preferably be measured as biomass to ensure consistency with D4C2. Where several stocks or assessed populations of a species reside within the foodweb assessment area, these stocks or assessed populations are treated as separate entities in the size distribution analysis.

Note that if only one species is present in the guild and that this species is already assessed under D1 or D3, no new assessment is made here. Where information only exists for some of the species included in the guild in D4C2, the guild can still be assessed as long as the species included are considered representative of the guild.

Indicators for D4C4

Examples of indicators of D4C4 include primary productivity, mean recruitment success, mean somatic growth, mean condition factor, and mean total mortality within a guild (ICES, 2014; Shephard *et al*., 2014, Eero *et al*., 2015). Several of these aspects are highly spatially and seasonally variable, and the indicators should therefore be estimated at the same time of year and with a fixed spatial coverage for a given species. Other indicators of guild productivity may also be appropriate.

Productivity is highly influential in biomass development, and D4C4 is therefore strongly linked to D4C2. It can in many cases be considered an early warning indicator for change in D4C2. For this link to be direct, the species included under this criteria in a guild should be representative of species included under D4C2. Note that if only one species is present in the guild and this species is already assessed under D1 or D3, no new assessment is made here.

Implications of the D4 assessment for other descriptors

Observed ecosystem and foodweb changes potentially provide information that can be used to improve the predictions of future stock productivity under D3. For example, an increased population of seals is likely to result in a sustained elevated natural mortality of its prey. Changes to the abundance and composition of zooplankton, as often anticipated under climate change, is likely to impact on the productivity of planktivorous fish, including the larval stages of most of the commercially exploited stocks. Species that are introduced or are migrating into a new area may increase in abundance to the level where they significantly alter the foodweb, either increasing predation, decreasing food abundance, or providing a new alternative food source. All of these provide easy to understand examples of interactions between assessment elements of descriptors 1, 2, and 3, mediated through foodwebs as assessed under D4.

Sources and references

Beaugrand, G., Ibañez, F., and Reid, P. C. (2000). Spatial, seasonal and long-term fluctuations of plankton in relation to hydroclimatic features in the English Channel, Celtic Sea and Bay of Biscay. Marine Ecology Progress Series, 200, 93–102.

Boschetti, S., Piroddi, C., Druon, J., and Palialexis, A. 2021. Marine Strategy Framework Directive – Review and analysis of Member States' 2018 reports – Descriptor 4: Food webs, EUR 30652 EN, Publications Office of the European Union, Luxembourg, 2021, ISBN 978-92-76-32461-4. JRC124263. http://dx.doi.org/10.2760/32522.

Eero, M., Hjelm, J., Behrens, J., Buchmann, K., Cardinale, M., Casini, M., Gasyukov, P., et al. 2015. Eastern Baltic cod in distress: biological changes and challenges for stock assessment. ICES Journal of Marine Science, 72(8), 2180–2186. https://doi.org/10.1093/icesjms/fsv109.

EU. 2017. Commission Decision (EU) 2017/848 of 17 May 2017 laying down criteria and methodological standards on good environmental status of marine waters and specifications and standardized methods for monitoring and assessment, and repealing Decision 2010/477/EU (Text with EEA relevance). Official Journal of the European Union, L 125: 43–74. http://data.europa.eu/eli/dec/2017/848/oj.

HELCOM. 2021. HELCOM Workshop on Status Assessment of Commercial Fish (ComFish WS 1-2021). 28–29 August 2021. Online meeting. *Availabe at https://portal.helcom.fi/meetings/ComFish%20WS%201-2021-934/MeetingDocuments/Notes%20of%20ComFish%20WS%201-2021.pdf*.

ICES. 2014. Report of the Workshop to develop recommendations for potentially useful Food Web Indicators (WKFooWI), 31 March–3 April 2014, ICES Headquarters, Copenhagen, Denmark. ICES CM 2014\ACOM:48. 75 pp.

ICES. 2015. Report of the Workshop on guidance for the review of MSFD decision descriptor 4 – foodwebs II (WKGMSFDD4-II), 24-25 February 2015, ICES Headquarters, Denmark. ICES CM 2015\ACOM:49. 52 pp.

ICES. 2020. EU request for advice on developing appropriate lists for Descriptor 3, commercially exploited fish and shellfish, for reporting by EU Member States under MSFD Article 17 in 2024. *In* Report of the ICES Advisory Committee, 2020. ICES Advice 2020, sr.2020.13. https://doi.org/10.17895/ices.advice.7646.

ICES. 2020. Joint OSPAR/HELCOM/ICES Working Group on Seabirds (JWGBIRD; outputs from 2019 meeting). ICES Scientific Reports. 2:80. 101 pp. http://doi.org/10.17895/ices.pub.7466.

Olin, A. B., Banas, N. S., Wright, P. J., Heath, M. R., and Nager, R. G. 2020. Spatial synchrony of breeding success in the blacklegged kittiwake *Rissa tridactyla* reflects the spatial dynamics of its sandeel prey. Marine Ecology Progress Series, 638, 177–190. https://doi.org/10.3354/meps13252.

Peschko, V., Müller, S., Schwemmer, P., Mercker, M., Lienau, P., Rosenberger, T., Sundermeyer, J., and Garthe, S. 2020. Wide dispersal of recently weaned grey seal pups in the Southern North Sea. ICES Journal of Marine Science, 77(5), 1762–1771. https://doi.org/10.1093/icesjms/fsaa045.

Shephard, S., Rindorf, A., Dickey-Collas, M., Hintzen, N. T., Farnsworth, K., and Reid, D. G. 2014. Assessing the state of pelagic fish communities within an ecosystem approach and the European Marine Strategy Framework Directive. ICES Journal of Marine Science, 71(7), 1572–1585. https://doi.org/10.1093/icesjms/fsu005.

Vasilakopoulos, P., Konrad, C., Boschetti, S. T, and Palialexis, A. 2021. Marine Strategy Framework Directive, Review and analysis of Member States' 2018 reports. Descriptor 3: Commercial species, EUR 30660 EN, Publications Office of the European Union, Luxembourg, 2021, ISBN 978-92-79-34175-8. JRC124746. http://dx.doi.org/10.2760/40557.

Recommended citation: ICES. 2021. EU Technical Service on MSFD Article 8 guidance on undertaking assessments for Descriptor 3 (commercially exploited fish and shellfish) and Descriptor 4 (marine foodwebs). *In* Report of the ICES Advisory Committee, 2021. ICES Advice 2021, sr.2021.14, https://doi.org/10.17895/ices.advice.8817.

Annex 1 Article 8 guidance for undertaking assessments for Descriptor 3 (commercially exploited fish and shellfish)

Table A1.1 Descriptor 3 on commercially exploited fish and shellfish.

DESCRIPTOR PROFILE		D3C1 Fishing mortality (F)	D3C2 Spawning stock biomass (SSB)	D3C3 Age and size distribution
Primary criterion		X	X	X
Type (pressure, impact, state)		Impact	State	State
Annex III MSFD	State	Species/stock	Species/stock	Species/stock
	Pressure	Extraction of or mortality/injury to wild species including target and non-target species		
	Activity	Fish and shellfish harvesting (professional and recreational)		
Action level	Elements	(Sub)regional	(Sub)regional	(Sub)regional
	Threshold values	EU	EU	(Sub)regional
	Use of criteria	EU		
Criteria linkages		D1C1	D1C2, D4C1, D4C2	D1C3, D4C3
Descriptor linkages		D1, D4		

A) Elements

Definition of commercially exploited species

The regional list of commercial species is defined by identifying the minimum number of species/stocks that together contribute a regionally agreed percentage of reported regional (or subregional) landings by weight. The estimation should be based on a regionally (or subregionally) coherent collection of fisheries dependent data representing a time period of at least six recent years. The agreed percentage should not be less than 90%. Additionally, species/stocks that contribute to a regionally agreed percentage of reported regional landings by value can be added.

The list of commercially exploited species/stocks is identified at regional level unless a subregional level is deemed more ecologically relevant for the region. Individual Member States can choose to add species/stocks of national relevance to the regionally agreed upon list of commercially exploited stocks/species.

The selected species/stocks should remain on the list even if not contributing to the agreed percentage in subsequent MSFD assessments. However, if it can be clearly demonstrated that a stock has consistently achieved good status and the lack of contribution to the agreed percentage is caused by the species/stock no longer being fished at noticeable levels, the species can, after regional agreement, be removed from the list. Species/stocks which have been commercially important in previous MSFD assessments but are currently at very low abundance due to previous fishing impact should be retained on the regional list of commercially exploited species/stocks.

Species/stocks which are both on the list of commercially exploited species/stocks and on the list of fish species for the purpose of biodiversity assessment under D1 should be integrated under D1 using the assessment status determined under D3. The assessments of criteria D3C1, D3C2, and D3C3 will then be used to assess criteria D1C1 (together with other pressures), D1C2, and D1C3, respectively and integrated using methods agreed under D1. If a species included under D1 is divided into two or more stocks within the MSFD assessment area, only part of these stocks may be included in the list of commercially exploited species assessed under D3; the remaining stocks should be assessed under D1.

Data used and appropriate indicators

D3 status is assessed based on stock assessment results from ICES (Northeast Atlantic), STECF, and GFCM (Mediterranean and Black seas) wherever these are available. Stocks that are not assessed by ICES, STECF, or GFCM can be assessed by Member States following the principles of stock assessment outlined by any of the three bodies as well as the guidance given in this document. Data should be collected according to agreed standards for MSFD regions. The data for a stock to be assessed must as a minimum include information on either annual total catch as well as information on length distribution of catches, or annual total catch and time-series of biomass/abundance indices.

B) Assessment areas and scales

The status of individual commercial stocks/species should be evaluated and reported at the level of the stock, whereas the integrated D3 assessment should be conducted at the spatial scale most relevant to the majority of the stocks/species in each spatial assessment unit, which may correspond to an MSFD region, combined regions, a subregion, or (if relevant) a smaller unit. In cases where stocks/species occur in wider areas, they should be assessed in all assessment units in which they occur and are included on the regionally agreed list. This means that widespread stocks may contribute to the D3 status assessment in more than one assessment unit, as well as region.

C) Temporal aspect of assessment

To reflect the status of the descriptor within the six-year assessment period, indicator status in each of the six years should be considered in the evaluation. Temporal aggregation of the annual results are explained under point e) for each criterion.
As results of D3 are to be integrated with results under D1 and D4, the calendar years of the applied six-year assessment period should be coordinated across the three descriptors to ensure consistency in the evaluations of foodweb and diversity aspects.

D) Spatial aggregation of assessments

[No guidance needed as assessment results are provided by stocks]

E) Threshold values

GES and threshold values relate to approaches defined by ICES or GFCM for C1 and C2 and (sub)regionally agreed values for C3.

Criterion D3C1

The indicator used to assess status under D3C1 is the fishing mortality rate of the stock, and this indicator must be at or below the fishing mortality which in the long term leads to maximum sustainable yield (MSY), F_{MSY}. Alternative thresholds to F_{MSY} can be used in cases where a full management strategy evaluation (as conducted in support of e.g. the CFP or a multiannual plan [MAP]) has shown that these are more appropriate with respect to obtaining MSY and maintaining full stock productivity potential.

The CFP and associated MAPs refer to F_{MSY} as a target. The use of F_{MSY} as a target means that in an unbiased management system, where the prediction and implementation of F is almost perfect, the threshold F_{MSY} will be slightly exceeded in around half of the years. As a further complication, the MAPs also define F_{MSY} ranges of which the upper limit can be used if the stock is above the agreed precautionary biomass reference level and certain criteria are fulfilled. Annual fishing mortalities above F_{MSY} are hence in accordance with the MAPs in certain cases. However, MAPs also state that the plan aims to 'ensure that the conditions described in Descriptor 3 contained in Annex I to Directive 2008/56/EC[3] are fulfilled (and other available MAPs[4]).

To support alignment between the MFSD and the CFP, the average F over the assessment period for the MSFD (six years) should not exceed the F_{MSY} threshold, though this may occur in individual years as a result of the conditions laid out in the CFP and MAP.

For exploited stocks which remain below levels at which recruitment is considered impaired (B_{lim}), measures should be taken to reduce F (see e.g. MAPs/CFP). To fulfil the requirement for reduced pressure, the F threshold for such stocks must be lower than F_{MSY}. In such cases, the F threshold could be defined as linearly decreasing below MSY $B_{trigger}$. Hence, if biomass on average has been 80% of MSY $B_{trigger}$ in the assessment period, the F threshold should be 80% of F_{MSY}.

[3] (https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R1139&from=en.
[4] https://ec.europa.eu/oceans-and-fisheries/fisheries/rules/multiannual-plans_en.

F_{MSY} estimation is not simple for stocks which do not have full age- or biomass-based assessments. For these stocks, a variety of proxies are available depending on the existing information (e.g. Chong *et al*., 2020; Arkhipkin *et al*., 2021; Armelloni *et al*., 2021; ICES, 2021a). A recommendation of methodological approaches should be addressed in a future update of this guidance document.

Criterion D3C2

The indicator used to assess status under D3C2 is the spawning-stock biomass (SSB) of the stock, and this indicator must be above biomass levels capable of producing MSY. The CFP and the associated MAPs refer to biomass thresholds in accordance with ICES MSY approach, something that is also being applied in the Mediterranean context (see e.g. MAP in the Western Mediterranean Sea; EU, 2019). ICES MSY approach defines the biomass capable of producing MSY as the 5th percentile of the projected annual biomass of a stock fished for long time periods at F_{MSY} (MSY $B_{trigger}$) with the added requirement that the MSY $B_{trigger}$ cannot be less than the biomass required to ensure that recruitment is impaired (below B_{lim}) in less than 5% of the years in the long-term projection (B_{pa}; ICES, 2021a), i.e. $B_{threshold}$ = max(B_{pa}, MSY $B_{trigger}$). Alternative approaches to identify thresholds can be used in cases where a full management strategy evaluation (as conducted in support of e.g. the CFP or a MAP) has shown that these are more appropriate with respect to obtaining MSY and maintaining full stock productivity potential. In the absence of information on B_{lim}, B_{pa}, and MSY $B_{trigger}$, regionally agreed proxies of unfished biomass can be used as the basis for estimating the biomass threshold (e.g. B_{lim} = 0.2*B_{0proxy}, B_{pa} = 2*$B_{lim;}$; see also the approach from FAO-GFCM [2014]).

For short-lived species, the biomass threshold is defined in relation to a target, with B_{pa} representing a level to ensure MSY with a 95% probability of not adversely affecting recruitment (ICES, 2021a). For other stocks, the biomass threshold $B_{threshold}$ sometimes represents a limit that should not be subceded on average in simulations to determine F_{MSY}. These considerations would also apply for stocks if there was perfect understanding of stock dynamics and perfectly implemented management. To accommodate this interannual variability, the average biomass over the assessment period for the MSFD (six years) for all stocks (short-lived or otherwise) should not be less than the biomass threshold though this may occur in individual years as a result of the conditions laid out in the CFP and MAPs.

For stocks which do not have full age- or biomass-based assessments, $B_{threshold}$ estimation is not simple. For these stocks a variety of proxies can be used depending on the information available (Chong *et al*., 2020; Arkhipkin *et al*., 2021; Armelloni *et al*., 2021; FAO-WGSAD, 2021; ICES, 2021a), including but not limited to biomass-related indices such as catch per unit effort or survey abundance indices. Results should be reported in a way that reflects on potential biases and limits to interpretation with respect to the data used as well as confidence in the methods, and threshold values should be provided. Examples of some application at regional level may be found in the Mediterranean Ecological Quality Status Report (UNEP, 2017).

Certain stocks/species which are severely depleted compared to historical levels may not have sufficient information for a full analytical assessment with associated thresholds. Rather than reporting such stocks as having 'non assessed' status, they could be evaluated as having 'biomass depleted below the last estimated reference level' with an associated estimated confidence in that rating based on the data and information available. For example, if survey catches of the species have not increased, it is unlikely that the stock has rebuilt.

Species which have historically been lightly fished, and for which biomass indices from surveys can be assumed to be indicative of unfished biomass for at least some years of the time-series, are typically not the target of directed fisheries. As bycatch species, they may be more or less sensitive to fisheries. Species which are not sensitive (ICES, 2021b) and which are only lightly impacted by fisheries can possibly be evaluated by regional expert judgement (far from likely reference points; close to likely reference points) provided that the results are reported with a clear indication of the quality of this information. In case of doubt, the evaluation should be reported as not assessed.

Criterion D3C3

Approaches to define threshold values for D3C3 indicators have yet to be agreed. Further, compatibility between any threshold values of D3C3 and the threshold values of criteria D3C1 and D3C2 should be ensured to ascertain that all criteria can in fact be attained simultaneously.

The status should be evaluated using the average over the assessment period for the MSFD (six years) as status in individual years may fall on either side of the threshold due to natural variability, and it should be considered that both very high and very low values might be undesirable for certain indicators.

Comparisons of D1 and D3 show that the formulations of criteria D1C3 and D3C3 differ in their contents, with D3C3 being considerably more restrictive than D1C3 as it pertains only to changes in stock size and age structure whereas D1C3 also pertains to species productivity. As several commercial stocks have exhibited changes in stock productivity due to e.g. changes in individual growth, condition, maturity, recruitment, and mortality, it is recommended to expand the focus of D3C3 beyond size- and age-composition and genetic diversity. In this regard, criteria D3C3 can function as an early warning indicator for D3C2 and thereby D3C1.

Potential indicators for the current D3C3 include:

- Length distribution. Suggested indicator: median, 95% or other quantiles of length of individuals in the population. This indicator will reflect the combined impacts of recruitment, individual growth, and mortality. It should preferably be monitored by fish length distribution from stock assessments combining length distribution-at-age and numbers-at-age in the population, or data on length distribution from surveys. Where representative survey catches are lacking, data on length distribution in catches may be used. Results should be reported in a way that reflects on potential biases and limits to interpretation with respect to the data used.
- Age distribution. Suggested indicator: median, 95% or other quantiles of age of individuals in the population. This indicator will reflect the combined impacts of recruitment and mortality. It should be monitored by age distribution from stock assessment estimates of age and numbers-at-age in the population, or data on age distribution from surveys. Where representative survey catches are lacking, data on age distribution in catches may be used.
- Maturity. Suggested indicator: length at 50% maturity or age at 50% maturity. Maturity should be estimated from representative sampling of the population in a biologically relevant season. Smaller length- or age-at-maturity can be indicative of high mortality, changes in size selective mortality, or of both enhanced or deteriorating individual condition. It should be noted that a decreasing age at 50% maturity can be a sign of serious stress in the stock, and observations on this should therefore not lead to increasing thresholds for D3C1 or increasing estimates of SSB in D3C2.

Potential indicators of an expanded D3C3 include:

- Recruitment. Suggested indicator: recruitment per spawner. It should be estimated by SSB and the number in the population at the earliest observed age from stock assessments, or by data on recruitment and SSB from surveys. Where representative survey catches are lacking, data on recruitment and SSB from catches may be used. For stocks with high fishing impact on juveniles, recruitment can be measured at later ages and will then incorporate the juvenile mortality induced by fishing.
- Individual growth. Suggested indicator: mean weight-at age-anomaly averaged across appropriate ages. It should be estimated by weight-at-age in the stock from surveys or, where representative survey catches are lacking, data from catches.
- Condition. Suggested indicator: mean condition or mean relative condition (ratio between observed weight-at-length and predicted weight-at-length based on a reference data set for the stock) averaged across individuals or proportion of individuals in poor condition (poor condition to be defined on a stock basis). It should be estimated from condition in the stock from surveys.
- Natural mortality. Suggested indicator: estimated natural mortality. Increased natural mortality due to e.g. predation or disease can decrease stock productivity to levels where previously estimated thresholds of fishing mortality are no longer sustainable and vice versa. With the rebuilding of apex predator populations, the natural mortality of particularly planktivores and sub-apex predators is likely to increase. Natural mortality can be estimated from multispecies models. However, these models are not widely available and difficult to validate.

Use of criteria:

To ensure consistency with existing management approaches, the following integration rules are suggested with respect to the potential combination of assessment results for D3C1 and D3C2. In the absence of agreed methodologies and thresholds, D3C3 is not included in the current integration rules, but is assessed qualitatively using indicative threshold values or descriptions of trends. The use of D3C3 will be addressed in the next update of this guidance:

Table A1.2 Suggestions for D3C1 and D3C2.

D3C1	**D3C2**	**Stock status**
GES	GES	GES
GES	Not in GES	Not in GES
Not in GES	GES	Not in GES
Not assessed; short-lived stocks managed by $B_{escapement}$	GES	GES
Not assessed; short-lived stocks managed by $B_{escapement}$	Not in GES	Not in GES
Not assessed; all other stocks than short-lived stocks ones managed by $B_{escapement}$	GES	Not assessed
Not assessed; all other stocks than short-lived stocks managed by $B_{escapement}$	Not in GES	Not in GES
GES	Not assessed	Not assessed
Not in GES	Not assessed	Not in GES

Integration across stocks is not conducted. Illustration of the assessments are addressed in Section G).

F) Confidence

The D3 assessment should be accompanied by an evaluation of the confidence in the assessment per stock. The evaluation of confidence should include (per criterion/parameter addressed): the appropriateness of stock data with respect to length of time-series and spatial coverage of data collected, the appropriateness of the models used for estimating indicators, and the appropriateness of the approach for estimating stock thresholds values. Further guidelines for estimating confidence need to be developed.

G) Presentation of assessment results

D3 criteria assessments are presented using bar charts depicting the proportion/number of stocks in GES, not in GES, and unassessed in the agreed assessment area (super-regional, regional, subregional, or, if relevant, smaller). An example illustration is given in Figure A1.

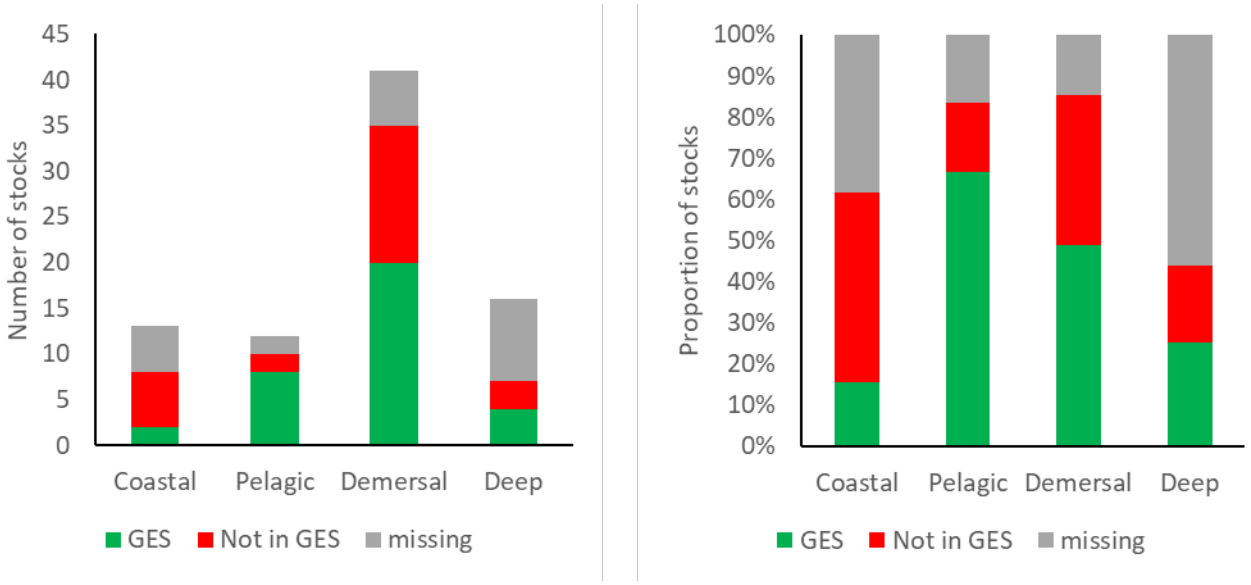

Figure A1.1 Examples illustrating D3 status by stock category (based on criteria D3C1 and D3C2). Red: below GES, green: GES, grey: missing (not assessed).

H) Gaps in knowledge and outstanding issues

Further development of indicators and thresholds for D3C3

Further discussion on the potential need for broadening the indicators under D3C3 to reflect changes in productivity beyond size structure and evaluation of the methodological approaches should be conducted. This work should provide firmer guidelines on which indicators and approaches are most appropriate for specific regions and stocks and how thresholds should be defined.

Changes in stock productivity and comparability of MSFD assessments over time

Stock productivity and hence F_{MSY} varies over time. To adjust for this, F_{MSY} is regularly updated to match current productivity conditions. Such updated F_{MSY} pertains to current conditions but not historical conditions and using the value for a historical time period provides a biased view with respect to the development of fishing pressure relative to F_{MSY}. If F_{MSY} was increased with the update, stocks may appear historically underexploited and vice versa. Other updates are also frequently made, meaning it is difficult to derive historically comparable thresholds for F_{MSY} from the literature. In order to allow an unbiased assessment of the temporal development, F_{MSY} should ideally be re-evaluated for each six-year assessment period.

In some cases, predation from species monitored under D1 may impose significant mortality on the stock, and this link between D1, D4, and D3 should ideally be made clear in the reporting.

Possible managment implications of the D3C1 and D3C2 assessments

There are two major reasons why thresholds are not met on average over the assessment period: the governance system does not implement and control sufficiently restrictive fishing opportunities, and the scientific advisory system provides inaccurate advice. Failure related to the governance system may reflect either e.g. a lack of international agreement on shared stocks or inefficient implementation of catch and effort limits within the EU (e.g. through a non-enforced limit to fishing effort or catches). In these cases, giving the stock a red status can create awareness and pressure for greater adherence to scientific advice. For stocks where F regulation is not effective, other measures should be taken to ensure that sufficient individuals are allowed to spawn to maintain the stock, e.g. implementation of minimum (or maximum) conservation reference size or areas closed to fishing.
The scientific advisory system may provide inaccurate advice if changes to stock productivity lead to biased stock assessments, which in turn lead to too high or too low catch advice. This happens if the stock has recently experienced greater than expected mortality or lower than expected growth and recruitment, situations currently observed in several stocks. If these situations persist, they often leads to a downscaling of F_{MSY} , which results in assessments that may indicate that the F threshold has been exceeded in previous years, where a higher target F was used in advice, even in cases where governance actions may have been appropriate and sufficient according to the prevailing scientific advice. The updated assessment will show that the stock was overexploited in spite of the governance actions being appropriate and sufficient at the time of implementation. It would be informative to include information on the likely role of each of these factors along with the confidence in the assessment.

Assessment of stocks without analytical stock assessments and associated thresholds

Stocks which occur in the regional list of commercial species but do not have a full analytical assessment with associated reference points may not have historically been considered a priority for assessment. Such data deficient stocks may include both lightly impacted as well as heavily impacted stocks which are at too low levels to allow sufficient data to be collected. These gaps may potentially open opportunities for commercial (or overfished) stocks to be excluded from the D3 assessments due to data shortage. With their appearance on the D3 list of commercial stocks/species, data should be assembled at the appropriate spatial level and attempts made to construct an assessment model which can be used to assess the status of the stock relative to MSY reference levels. If it is not possible to conduct an analytical assessment with associated reference levels, data may allow the use of dedicated methods based on e.g. indices of abundance/biomass, length distribution, and catches or other data-limited approaches.

Stock selection and the definition of methods to be applied should be agreed at regional level. However, in the absence of such agreements, the lack of established thresholds should not impede Member States from reporting on stock status. In this case, taking into account data quality and quantity, assessment should be carried out prioritizing the use of most adequate international standards while minimizing uncertainties associated with the assessment.

References

Arkhipkin, A. I., Hendrickson, L. C., Payá, I., Pierce, G. J., Roa-Ureta, R. H., Robin, J. P., and Winter, A. 2021. Stock assessment and management of cephalopods: advances and challenges for short-lived fishery resources. ICES Journal of Marine Science, 78(2), 714–730. https://doi.org/10.1093/icesjms/fsaa038.

Armelloni, E. N., Scanu, M., Masnadi, F., Coro, G., Angelini, S. and Scarcella, G. 2021. Data Poor Approach for the Assessment of the Main Target Species of Rapido Trawl Fishery in Adriatic Sea. Front. Mar. Sci. 8:552076. https://doi.org/10.3389/fmars.2021.552076.

Chong, L., Mildenberger, T. K., Rudd, M. B., Taylor, M. H., Cope, J. M., Branch, T. A., Wolff, M., and Stäbler, M. 2020. Performance evaluation of data-limited, length-based stock assessment methods. ICES Journal of Marine Science, 77(1), 97 –108. https://doi.org/10.1093/icesjms/fsz212.

EU. 2019. Regulation (EU) 2019/1022 of the European Parliament and of the Council of 20 June 2019 establishing a multiannual plan for the fisheries exploiting demersal stocks in the western Mediterranean Sea and amending Regulation (EU) No 508/2014. https://eur-lex.europa.eu/eli/reg/2019/1022/oj.

FAO-GFCM. 2014. Report of the fifteenth session of the SAC Subcommittee on Stock Assessment (SCSA). Bar (Montenegro). https://gfcmsitestorage.blob.core.windows.net/documents/Reports/GFCM-Report-2014-SCSA-15.pdf.

FAO-WGSAD. 2021. Report of the Working Group on Stock Assessment of Demersal Species (WGSAD). https://gfcm.sharepoint.com/EG/Report%20v2/Forms/AllItems.aspx?id=%2FEG%2FReport%20v2%2F2021%2FWGSAD%2FWGSAD%5F2021%5FReport%2Epdf&parent=%2FEG%2FReport%20v2%2F2021%2FWGSAD&p=true.

ICES. 2021a. Tenth Workshop on the Development of Quantitative Assessment Methodologies based on LIFE-history traits, exploitation characteristics, and other relevant parameters for data-limited stocks (WKLIFE X). ICES Scientific Reports, 2:98. 72 p. http://doi.org/10.17895/ices.pub.5985

ICES. 2021b. Workshop on Fish of Conservation and Bycatch Relevance (WKCOFIBYC). ICES Scientific Reports, 3:57. 125 pp. https://doi.org/10.17895/ices.pub.8194.

UNEP. 2017. The First Quality Status Report for the Mediterranean – MED QSR (2017). https://www.medqsr.org/sites/default/files/inline-files/2017MedQSR_Online_0.pdf.

Annex 2 Article 8 guidance for undertaking assessments for Descriptor 4 (marine foodwebs)

Table A2.1 Descriptor 4 on foodwebs.

DESCRIPTOR PROFILE		D4C1 Diversity of the trophic guild	D4C2 Balance of abundance between trophic guilds	D4C3 Size distribution of individuals across the trophic guild	D4C4 Productivity of the trophic guild
Primary criterion		X	X		
Type (pressure, impact, state)		State	State	State	State
Annex III MSF	State				
	Pressure	Extraction of or mortality/injury to wild species including target and non-target species			
	Activity	Fish and shellfish harvesting (professional and recreational)			
Action level	Elements	(Sub)regional	(Sub)regional	(Sub)regional	(Sub)regional
	Threshold values	(Sub)regional	(Sub)regional	(Sub)regional	(Sub)regional
	Use of criteria	(Sub)regional			
Criteria linkages		D1C2, D3C2	D1C2, D1C6, D2C2, D3C2, D5C2	D1C3, D3C3	D1C2, D1C3, D3C2, D3C3
Descriptor linkages		D1, D2, D3, D5, D6			

A) Elements

Trophic guild definition

Table A2.2 Suggested trophic guilds to enhance regional comparability.

Guild	Description	Example species/groups
Pelagic primary producers	Phytoplankton	Diatoms, dinoflagellates
Benthic primary producers	Macrovegetation, included where relevant for the foodweb assessment area	Seagrass meadows, kelp forests
Secondary producers	Mesozooplankton (200 micron–20 mm)	Copepods, cladocerans
Benthic filter-feeding invertebrates	Benthic filter-feeding invertebrates, included where relevant for the foodweb assessment area	Mussel, scallops, brittle stars
Benthic feeding invertebrates	Benthic invertebrates feeding predominantly on detritus or other benthic invertebrates, and/or constituting prey for sub-apex predators.	*Nephrops*, crabs, shrimps
Planktivorous fish and invertebrates	Fish and invertebrates feeding predominantly on zooplankton	Anchovy, herring, horse mackerel, jellyfish
Sub-apex pelagic predators	Fish and invertebrates feeding pelagically on fish and other prey types	Mackerel, saithe, tunids, Loligo
Sub-apex demersal predators	Fish and invertebrates feeding demersally or on the bottom on fish and other food	Sole, hake, haddock, octopus
Mammal apex predators	Marine mammal piscivores feeding on sub-apex predators	Killer whale
Fish apex predators	Fish piscivores feeding on sub-apex predators	Large tuna, large cod, large sharks

The suggested guilds could be further subdivided into sub-guilds (e.g. sub-apex demersal predators feeding on smaller items and sub-apex demersal predators feeding on larger prey items) where relevant for the assessment area and agreed regionally.

Assigning individuals to guilds in a foodweb assessment area

The guild concept is broader than a simple estimation of trophic level by species. However, estimates of trophic level may help to assign species to guilds. Estimated trophic levels can be derived from a variety of methods including the literature, diet analyses, ecosystem models, stable isotope analyses, and genetic analyses. Species could also be

assigned to guilds based on diet composition at a given size, although this requires more data. Estimated diet composition can vary between locations for the same species, as well as between different assessments and years. Hence, estimations may need to be repeated to ascertain whether diet compositions change over time.

A species often belongs to different guilds through its life cycle (e.g. planktivorous, benthivorous, and finally piscivorous). With the rather broad guild categories given in Table A2.2, most species will remain in a guild for the duration of their life after the early stages (e.g. flatfish will remain in "sub-apex demersal predators" and herring in "planktivorous fish and invertebrates"). However, there are species that grow sufficiently large to achieve a diet with a large contribution of sub-apex predators, thus making them apex predators at large sizes. Such species can be allocated to different guilds at different ages or sizes.

Stocks or assessed populations straddling several foodweb assessment areas such as mackerel, tuna and larger whales, should contribute to a certain foodweb assessment area depending on the proportion of the stock residing in the area and the time spent in it. Invasive species should be included where they make an important contribution to the foodweb in terms of biomass.

Indicator estimation considerations with respect to guilds

If information only exists for some of the species included in the assessed guild, the guild can still be assessed as long as the species included are considered representative of the entire guild. All individuals contributing to a guild cannot be monitored in the marine environment. Instead, the guild-based indicators are based on the monitoring or modelling of representative components in the ecosystem. In cases where guilds are monitored through only a few of their components, it is particularly important to investigate whether these components are likely to be representative of the remainder of the guild.

Considering that data for very small individuals are often highly uncertain, adding data on these smaller individuals may add more noise than signal to the estimated foodweb indicators. In order to ensure that signals in data are not swamped by uncertainty, smaller individuals can be excluded from the indicator estimation assigning species to guilds based on diet from the life stages at which reliable data on e.g. size and biomass becomes available. This lower size or age cut-off should be constant through time.

B) Assessment areas and scales

Foodwebs should be assessed at regional level in regionally coordinated actions. Subregional assessment can be applied provided there is scientific evidence showing that this a more appropriate scale to assess the trophic guilds and their links to other descriptors. It is only appropriate to assess foodwebs in a Member State EEZ in rare cases. However, if a Member State collects information to suggest that smaller areas are consistent with these requirements, additional foodweb assessments can be completed for these smaller areas.

The spatial scale at which the foodweb is assessed should be selected to be relevant to most guilds. The foodweb assessment area should be chosen to either have limited migration/exchange or good information on the timing and extent of migration/exchange. Compatibility of results with those of other descriptors can be fulfilled by assessing foodwebs in a spatial area where there is limited net transport of biomass across the area boundary for the concerned guilds.

Based on its mean depth, each foodweb assessment area is assigned as either 'coastal'(mean depth 0–50 m), 'shelf' (50–200 m), 'slope' (200–800 m), or 'deep' (> 800 m).
It should be noted that while entirely pelagic foodwebs can be analysed where relevant, based on the selection of trophic guilds included in the assessment, they are still assigned to one of these four features according to the bottom depth.

C) Temporal aspect of assessment

To reflect the status of the descriptor within the six-year assessment period, indicator status in each of the six years should ideally be considered in the evaluation. As results of D4 are dependent on results under D1, D2, D3, D5, and D6,

the calendar years in this six-year assessment period should be coordinated across these descriptors. The assessment of D4 is based on the average indicator over the assessment period. The status assessment can be accompanied by a description of trends to facilitate the use of D4 as an early warning indicator.

D) Spatial aggregation of assessments

Spatial aggregation of assessments is not considered relevant for foodwebs except in rare cases.

E) Threshold values

Indicators for all four criteria under D4 are assessed separately for each guild.

As the indicators are considered highly impacted by natural variation and complex interactions, threshold values should be defined as ranges, with lower and upper levels. The threshold range should encompass natural variability and the possibility of the indicator to reflect structural changes over time. Indicator values inside the threshold range correspond to GES.
Data for a species/stock should be consistent with the estimated biomass within the foodweb assessment area under all D4 criteria, and the indicator assessment requires consideration of cross descriptor spatial coherence between D1, D2, D3, D5, and D6.

Indicators for D4C1: diversity of the guild

Indicators for this criterion reflect the relative abundance of species (also known as species dominance) or the relative change in abundance of species within the guild. Examples include diversity indices such as the Simpson or Shannon indices. Indicators should preferably measure abundances in terms of biomass to ensure consistency with D4C2. Indicators of the proportion of species that are increasing/decreasing (e.g. seabirds [OSPAR/HELCOM/ICES JWGBIRD, 2020]) can also be used if they reflect development in a guild and even if they do not reflect species diversity directly.

Data at the lowest available and applicable taxonomic level should be used. Where several stocks or assessed populations of a species reside within the foodweb assessment area, these stocks or assessed populations are treated as separate entities in the diversity analysis. Data representing several taxonomic levels can be included in the analysis as long as data for a specific stock/species only appears once (e.g. division between *Sepia elegans* and "all other Sepiidae").

Indicators for D4C1 are only applicable for guilds that contain at least 2 species/species groups. Where taxonomic information is not available for a guild, a D4C3 indicator based on size distribution should be used instead.

Indicators for D4C2: balance in abundance of guilds

This section addresses the estimation of biomass by guild. The balance between guilds is assessed in the subsequent integration of indicators across guilds.

All regionally relevant guilds should be included in D4C2 to allow for the subsequent assessment of the balance between guilds. The indicators used to assess D4C2 should provide estimates of development in the biomass of each guild, which should be scaled to provide information on total biomass of the guild in the foodweb assessment area. Additionally, information on the surface area (in km^2) of the foodweb assessment area should be included to facilitate comparisons with other assessment areas and (possibly) regions. Regional agreement on the way in which biomass for any non-assessed species or subareas is included is key to the assessment result.

The biomass estimates can be attained in various ways. Common for most approaches is that they use models (simple as well as complex) to derive biomass from samples, which usually represent temporal development. The applied approaches should be regionally agreed upon and quality assured to ascertain their capacity to estimate biomass.

Biomass of primary producers can, as a minimum acceptable indicator, be assessed using ocean greenness together with a transformation factor to transform this to biomass. Methods and indicators developed for D5 or D1 pelagic

habitats (plankton biomass) can be used either to supplement this or as standalone estimates where they are appropriate for the entire foodweb assessment area and annual phytoplankton development cycle.

Biomass of mesozooplankton (200 micron–20 mm) is ideally estimated from repeated annual wide scale sampling. This can, however, be difficult to attain due to low or inconsistent sampling effort. Member States are encouraged to initiate coordinated widespread spatiotemporal sampling of mesozooplankton to ensure that the large variability in e.g. spring blooms is covered and to make use of existing data whenever possible.

Biomass of benthic filter-feeding invertebrates should as a minimum include the biomass of benthic filter-feeding invertebrates assessed under D3, D5, or D6. Where species biomass data is considered unreliable or highly uncertain, it should not be included. The biomasses of species which are not assessed should be investigated to ideally determine the likely scale of their contribution to the guild biomass.

Biomass of benthic-feeding invertebrates should as a minimum include the biomass of assessed stocks of benthic-feeding invertebrate predators under D3, D5, or D6. Where species biomass data is considered unreliable or highly uncertain, it should not be included. The biomasses of species which are not assessed should be investigated to ideally determine the likely scale of their contribution to the guild biomass.

Biomass of planktivorous fish and invertebrates should as a minimum be assessed using all existing stock assessments of planktivorous fish and invertebrates. Information from surveys can be added for species which are not assessed, using methods such as those of Yang (1982) or Walker *et al*. (2017) to adjust for differences in catchability. The biomasses of species which are not assessed and not representatively caught in surveys should ideally be investigated to determine the likely scale of their contribution to the guild biomass. Where species biomass estimates are considered unreliable or highly uncertain, they should not be included in indicators under D4. This will generally include the large undersampled biomass of individuals < 5–10 cm. The exact cut-off level within this span will depend on regional monitoring methods but should remain constant over assessment years. Additional taxa that may be relevant in some assessment areas include filter-feeding animals such as baleen whales (D1) and jellyfish.

Biomass of sub-apex pelagic predators should as a minimum be assessed using all existing stock assessments of sub-apex pelagically feeding fish and invertebrates. Information from surveys can be added for species which are not assessed using methods such as those of Yang (1982) or Walker *et al*. (2017). The biomasses of species which are not assessed and not representatively caught in surveys should ideally be investigated to determine the likely scale of their contribution to the guild biomass. Where species biomass estimates are considered unreliable or highly uncertain, they should not be included in indicators under D4. This will generally include the large undersampled biomass of individuals < 5–10 cm. The exact cut-off level within this span will depend on regional monitoring methods but should remain constant over assessment years. Additional taxa that may be relevant in some assessment areas include pelagically feeding seabirds assessed under D1.

Biomass of sub-apex demersal predators should as a minimum be assessed using all existing stock assessments of sub-apex demersally feeding fish and invertebrates. Information from surveys can be added for species which are not assessed using methods such as those of Yang (1982) or Walker *et al*. (2017). The biomasses of species which are not assessed and not representatively caught in surveys should ideally be investigated to determine the likely scale of their contribution to the guild biomass. Where species biomass estimates are considered unreliable or highly uncertain, they should not be included in indicators under D4. This will generally include the large undersampled biomass of individuals < 5–10 cm. The exact cut-off level within this span will depend on regional monitoring methods but should remain constant over assessment years. Additional taxa that may be relevant in some assessment areas include demersally feeding seabirds assessed under D1.

Biomass of apex fish and invertebrate predators should as a minimum be assessed using all existing stock assessments of predators on sub-apex predators. As a minimum, the estimated biomass should use all existing stock assessments, adding information from surveys for species which are not assessed. Species that grow into being apex predators at a specific size or age can be included by agreeing regionally to split their biomass according to size or age, e.g. into sub-apex and apex predators. Note that if biomass of the species is included in e.g. the sub-apex predators, the same biomass should not also be included in apex predators.

Biomass of apex marine mammal predators should include the combined biomass estimates of mammals feeding on sub-apex predators. Examples include mammals monitored under D1. Where only abundances in numbers are available, information on mean weight of individuals can be used to derive biomass.

Indicators for D4C3: size distribution within guilds

A size distribution indicator within a guild reflects a species' population composition, growth, and total mortality. Therefore, indicators under D4C3 are not independent of the indicators under the other criteria.

Further, size-based indicators are reliant on information on size distribution of the community, and the estimated indicator is highly dependent on the size selectivity of the method used to derive the data. Data on length distribution of individuals within the guild can be derived from either scientific surveys or stock assessments. The method used to derive the input data must be consistent over time.

Potential data sources include catch rates from scientific surveys, commercial catches, commercial landings, and age- or size-based stock assessments (age-based assessments must be combined with information on length distribution-at-age). Among these, survey catch rates and stock assessments are considered most reliable, as commercial catch data will be based on a variety of fishing gears each with its own size selectivity. Commercial landings are not considered appropriate because of the unknown quantity of discards of particularly smaller fish.

Indicator estimation should be performed separately for each guild, and biomass measures should be used preferably to ensure consistency with D4C2. Examples of size-based indicators applied for fish and zooplankton include the large fish indicator (LFI), typical/median/mean/95th percentile of length, and mean maximum length (MML). Among these, median or 95th percentile length is preferred over mean length due to the greater stability of the median when applied to highly skewed data. The LFI relies on the estimation of the size of a 'large fish' to be estimated separately by guild and region, and the indicator is therefore not directly comparable between regions but may provide useful results within a region. Mean maximum length (MML) integrates aspects of species diversity and size structure (mean possible length in the guild) and has the advantage that it can be estimated without information on size distribution of individuals in the guild. It is, however, less responsive to changes in mean size in the guild than median length and if used, the source of the maximum length used in the study should be clearly given and maximum length should be constant over time within the region. Size indicators can be based on weight or length.

If only one species is present in the guild and this species is already assessed under D1 or D3, no new assessment is made under D4C3 to ensure consistency.

Indicators for D4C4 : productivity of the guild

Examples of indicators that can be used to assess the productivity of a guild include primary productivity, mean recruitment success, mean somatic growth, mean condition factor and mean total mortality within a guild (ICES, 2014; Shephard *et al*., 2014, Eero *et al*., 2015). Several of these aspects are highly variable spatially as well as over the course of the year, and the indicators should therefore be estimated at the same time of year and for a fixed spatial coverage. Other indicators of guild productivity may also be appropriate.

If only one species is present in the guild and this species is already assessed under D1 or D3, no new assessment is made under D4C4.

Productivity has a large influence on biomass development and any change in productivity is likely to be followed by a subsequent change in D4C2. Hence, D4C4 can provide early warning for changes in D4C2. For this link to be direct, the species included under this criteria in a guild should be representative of species included under D4C2.

Identification of appropriate thresholds

There are as yet no agreed methods to estimate thresholds for foodweb indicators. Further, their large natural variability means that the link to direct human pressure and hence direct management actions can be difficult to

identify. It has been suggested that foodweb indicators can be used in management as early warning indicators, eliciting either further study or precautionarity in management aimed at other descriptors.

Thresholds should be developed both for criteria and guild as well as at the integration level where balance between guilds is assessed.

E) Use of criteria

Methods to integrate the different guilds or monitor the balance of the foodweb within and across criteria are still not at a development stage where tested and peer-reviewed approaches are commonly agreed upon. This is a key issue that needs to be addressed before assessing the overall status of the foodweb from the monitored guilds.

The evaluation of the balance is highly dependent on the guilds selected for assessment. All assessed guilds should be included in the integration.

There have been suggestions to base integration on various weighting approaches, the most notable being the mean trophic level (MTL). The MTL indicator is interpretable as the average trophic level of a unit of biomass in the system and is essentially the sum of guild biomass times the guild trophic level divided by total biomass. However, it is unclear how the indicator relates to balance between guilds and the balance between pelagic and demersal guilds.

Further work should be conducted to develop integration methods reflecting the balance of the foodweb. However, the lack of agreed integration methods should not be used as an excuse to refrain from reporting the status of individual guilds.

F) Confidence

Confidence should be assessed at the level of individual guilds and criteria. The confidence assessment should as a minimum reflect the degree to which the assessed indicator(s) cover(s) the guild, its spatial and temporal relevance, and confidence in data sources and methods. Further work should be conducted to develop guidelines for confidence assessments of D4.

G) Presentation of assessment results

Assessment results for guilds should be visually presented using pie charts with each of the four criteria having a quarter (Figure A2). Quarter colour depicts the status of each criteria for the guild (clockwise, criteria 4.1, 4.2, 4.3, and 4.4). The pie size can potentially be scaled to reflect total guild biomass, although a minimum size will have to be applied to smaller guilds to ensure that they are readable. The figure is ordered according to the position of the guild in the water column and could be accompanied by pictures/icons of species representing the guild. Any guilds that are not applicable in the foodweb assessment area would be omitted, or they could be visualized by a white circle when needed to clarify if the lack of assessment results is explained by lack of assessment or lack of ecological relevance.

Figure A2.1 Example to visually represent the assessment results using pie charts. Red: below GES, green: GES, grey: not assessed.

H) Gaps in knowledge and outstanding issues

One requirement of the MSFD is to account for changes in climate or the natural environment in indicators and/or the setting of thresholds. The large recent changes in climate and ecosystems have made this a key issue to address in guidance or, as a minimum, in enhanced studies.

The development of agreed methods to monitor the balance of the foodweb is a key issue for advancing further assessment of foodweb status. Further work should be conducted to develop integration methods across guilds and criteria and to reflect the balance of the foodweb. Agreed methods to define thresholds should be derived.

Additionally, D4C4 productivity measures should be detailed and reviewed with the aim of changing this indicator from secondary to primary.

Cross-descriptor integration of results

Observed ecosystem and foodweb changes can provide information needed for improving the predictions of future stock productivity under D3. For example, changes to the abundance and composition of zooplankton, as often anticipated under climate change, are likely to impact the productivity of planktivorous fish, including the larval stages of most of the commercially exploited stocks. Another example could be when an increased population of marine

mammals is likely to results in a sustained elevated natural mortality of their prey. Further, species introduced or migrating into a new area may increase in abundance to the level where they significantly alter the foodweb, either increasing predation, decreasing food abundance, or providing a new alternative food source. All of these provide easy to understand examples of interactions between assessment elements of descriptors 1, 2, and 3, mediated through foodwebs as assessed under D4.

References

Eero, M., Hjelm, J., Behrens, J., Buchmann, K., Cardinale, M., Casini, M., Gasyukov, P., et al. 2015. Eastern Baltic cod in distress: biological changes and challenges for stock assessment. ICES Journal of Marine Science, 72(8), 2180–2186. https://doi.org/10.1093/icesjms/fsv109.

ICES. 2014. Report of the Workshop to develop recommendations for potentially useful Food Web Indicators (WKFooWI), 31 March–3 April 2014, ICES Headquarters, Copenhagen, Denmark. ICES CM 2014\ACOM:48. 75 pp.

ICES. 2020. Joint OSPAR/HELCOM/ICES Working Group on Seabirds (JWGBIRD; outputs from 2019 meeting). ICES Scientific Reports. 2:80. 101 pp. http://doi.org/10.17895/ices.pub.7466.

Shephard, S., Rindorf, A., Dickey-Collas, M., Hintzen, N. T., Farnsworth, K., and Reid, D. G. 2014. Assessing the state of pelagic fish communities within an ecosystem approach and the European Marine Strategy Framework Directive. ICES Journal of Marine Science, 71(7), 1572–1585. https://doi.org/10.1093/icesjms/fsu005.

Walker, N. D., Maxwell, D. L., Le Quesne, W. J., and Jennings, S. 2017. Estimating efficiency of survey and commercial trawl gears from comparisons of catch-ratios. ICES Journal of Marine Science, 74(5), 1448–1457. https://doi.org/10.1093/icesjms/fsw250.

Yang, J. 1982. An estimate of the fish biomass in the North Sea. ICES Journal of Marine Science, 40(2), 161–172. https://doi.org/10.1093/icesjms/40.2.161.