

CEMP Guideline for Pelagic Habitats thematic assessment integration method

OSPAR Agreement 2023-07¹

1 Introduction

It is challenging to distil a simple overarching message about the biodiversity status of pelagic habitats across large sea areas. Pelagic habitats are complex natural systems which experience high variability across multiple temporal scales, including daily (e.g. diel vertical migration), monthly (e.g. seasonality), and yearly (e.g. long-term trends in abundance / biomass). They also exhibit high spatial variability across relatively short distances, however, much of this patchiness is smoothed out when examined at the scale of large sea areas. This characteristic of pelagic habitats should make them particularly suitable for integrating biodiversity status at the regional scale.

The three Pelagic Habitats indicators (PH1/FW5, PH2, PH3) currently evaluated for OSPAR assessment all report on different components of ecosystem functioning. These separate ecosystem functions collectively contribute to the overall biodiversity status of a sea area. A “Good” status should represent healthy conditions, while a “Not good” status for any of these functions should act as an important warning sign.

The purpose of this document is to provide a framework for integrating the results of the three pelagic habitats indicators to facilitate the determination of overall biodiversity status at the scale of MSFD pelagic habitat types, and a higher integration to the OSPAR regional scale. The information provided by this integration can be used to better inform policy makers and aid decision making around the management of pelagic habitats across the OSPAR maritime area.

This guideline document contains a description of the methods used to integrate results across the three pelagic habitats indicators, and a similar methodology to integrate the overall level of confidence in those results. The integration methods described in this document are an extension of current hierarchical integration methods used to determine biodiversity status for the three OSPAR pelagic habitats indicators and were developed from a Joint Research Centre report on integration methods for MSFD biodiversity assessments (Dierschke et al. 2021).

Although this document describes a set of integration methods which use an example based on the specific results of the pelagic habitats indicator assessments for the OSPAR QSR 2023, it is also intended that these methods can also be applied to other thematic assessment as well as for future assessments.

¹ English only

2 Description of the integration method

2.1 Overview of the approach

For the assessment of biodiversity status for Pelagic Habitats, changes in state in OSPAR Regions II, III, and IV are currently evaluated through an analysis of three biodiversity indicators:

- PH1/FW5 - Changes in plankton communities
- PH2 - Changes in plankton biomass and / or abundance.
- PH3 - Changes in biodiversity index(s)

PH1/FW5 is an indicator of ecosystem function, which measures changes in the abundance of important plankton functional groups or “lifeforms”.

PH2 is a bulk indicator which examines changes in the abundance of copepods (the largest contributor to total zooplankton abundance), and changes in the biomass of phytoplankton (the main marine primary producers).

PH3 is a biodiversity indicator which examines changes in composition, species turnover, and dominance within phytoplankton and zooplankton communities. For the current assessment, PH3 was evaluated as a common indicator in the Celtic Seas (OSPAR Region III), and with the status of candidate indicator through pilot assessments in the Greater North Sea (OSPAR Region II), and in the Bay of Biscay and Iberian Coast (OSPAR Region IV; **Table 1**). While the results of the PH3 analysis have been included in the current Pelagic Habitats Thematic Assessment, they were not considered in the region-specific summary or confidence assessment for pelagic habitats in the Greater North Sea (OSPAR Region II) and in the Bay of Biscay and Iberian Coast (OSPAR Region IV), since PH3 was not a common indicator in those regions during the current assessment cycle.

Table 1. The status of how each indicator was considered or evaluated across the five OSPAR regions.

	PH1/FW5	PH2	PH3
Arctic Waters (Region I)	Not assessed	Not assessed	Not assessed
Greater North Sea (Region II)	Common	Common	Pilot assessment
Celtic Seas (Region III)	Common	Common	Common
Bay of Biscay and Iberian Coast (Region IV)	Common	Common	Pilot assessment
Wider Atlantic (Region V)	Not assessed	Not assessed	Not assessed

Pelagic habitats indicator assessments were conducted at the scale of distinct pelagic habitats or “assessment units”. Indicator results at the scale of assessment units were further integrated by pelagic habitat type within each OSPAR region. Please refer to *Section 2.2 Description of the assessment units being applied* in this document for a detailed description of assessment units and MSFD pelagic habitat types.

To assign a designation of biodiversity status to pelagic habitats based on the integration of indicator results we applied a semi-quantitative methodology described in McQuatters-Gollop et al. (2022), which was developed from the lessons learned from the previous OSPAR assessment (Intermediate Assessment 2017). In this case, the status of a habitat type can be designated as either “Good”, “Unknown”, “Not good”, or “Not assessed” (**Table 2**). Following the criteria outlined in this study, if a pelagic habitat has been assessed, its status should by default be considered as either “Unknown” or “Not Good”. At this stage it is not realistic to assign “Good” status to pelagic habitats due to the uncertainty around defining pristine baseline conditions. It is difficult to develop meaningful assessment thresholds for plankton and generally not possible to determine whether a particular state is desirable or undesirable, except under specific circumstances, such as cases of eutrophication. Following this logic, the status of assessed pelagic habitats should be considered “Unknown” by default. In cases when change has been detected and that change can be confidently linked to the impact of an anthropogenic pressure, the status of this habitat should then be considered “Detected trend linked to human activity (not good)”.

The current CEMP guidelines include an integration method to evaluate confidence in the results. The current confidence integration relies heavily on expert judgement and is therefore not entirely objective. For future assessments, it will be important to further develop the confidence assessment integration by applying a more quantitative approach (e.g. by applying integer values to evidence and agreement values and calculating a mean).

Table 2. Biodiversity status categories and colours used for the interpretation of indicator biodiversity state (adapted from McQuatters-Gollop et al. 2022).

Detected trend linked to human activity (not good)	Indicator value is below assessment threshold, or change in indicator represents a declining state, or indicator change is linked to increasing impact of anthropogenic pressures (including climate change), or indicator shows no change but state is considered unsatisfactory
Unknown	No assessment threshold and/or unclear if change represents declining or improving state, or indicator shows no change but uncertain if state represented is satisfactory
Detected trend not linked to human activity (good)	Indicator value is above assessment threshold, or indicator represents improving state, or indicator shows no change but state is satisfactory
Not assessed	Indicator was not assessed in a region due to lack of data, lack of expert resource, or lack of policy support.

2.2 Description of the assessment units being applied

Pelagic Habitats indicators were analysed to evaluate biological changes occurring for a set of assessment units and fixed-point stations across the Greater North Sea (OSPAR Region II), Celtic Seas (OSPAR Region III), and the Bay of Biscay and Iberian Coast (OSPAR Region IV). The "COMP4 assessment units" (Common Procedure for the Identification of the Eutrophication Status of the OSPAR Maritime Area, 4th application), an OSPAR data product, were used to spatially subdivide plankton samples (**Figure 1**). These assessment units are a geographical representation of the conditions most likely to drive plankton distribution, dynamics, and community composition (Enserink et al. 2019).

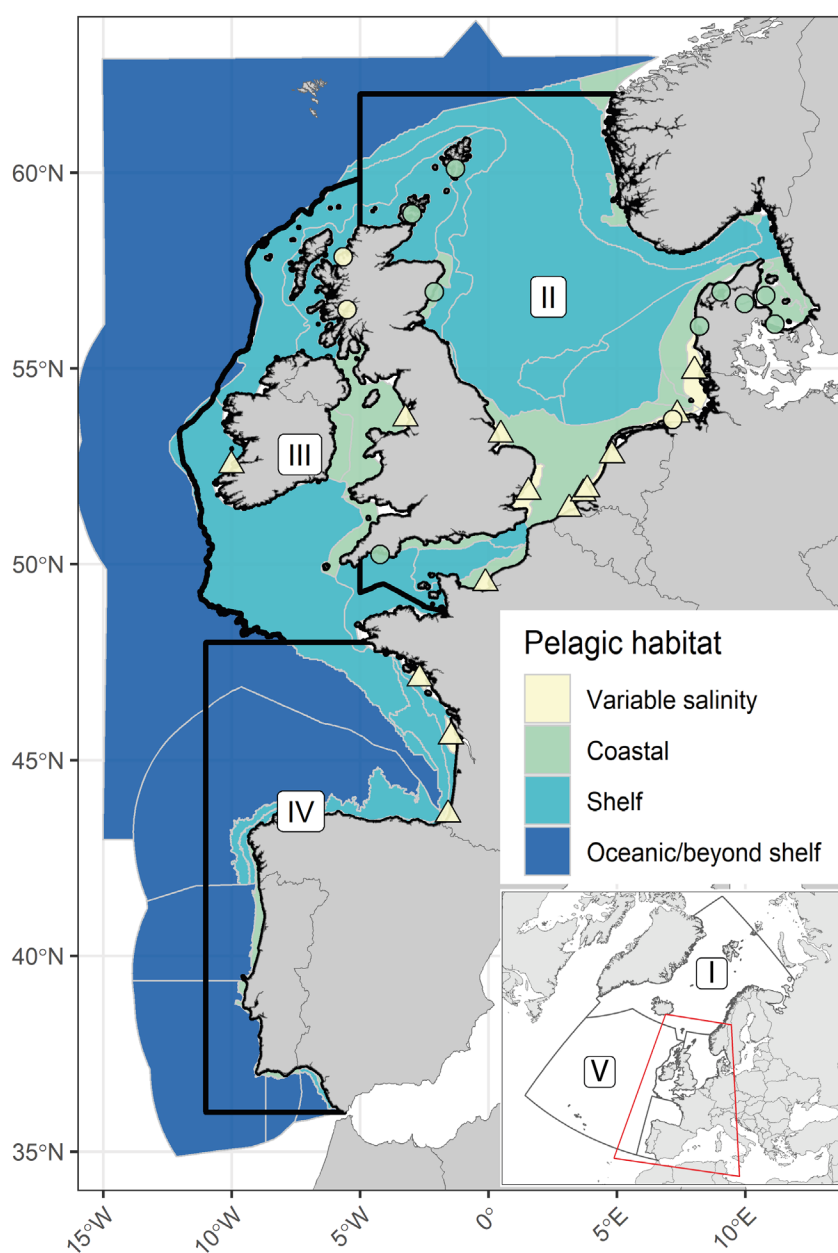


Figure 1. The distribution of the four pelagic habitat types across the three OSPAR Regions considered in the Pelagic Habitats indicator assessments, and the boundaries of the five OSPAR Regions across the OSPAR maritime area (inset). Fixed-point stations are represented as circles and river plumes are

represented as triangles. Boundaries between COMP4 assessment units used in this assessment are indicated in grey.

For OSPAR Regions II, III and IV, assessment units and fixed-point stations were grouped according to four pelagic habitat types so that indicator results could be integrated by habitat (**Figure 1**). The habitat type categories were created to align assessment outputs with EU MSFD features, with a view of allowing Contracting Parties that use the information for MSFD Art. 8 reporting. The four pelagic habitat types used in the assessment were:

- Variable salinity habitats
- Coastal habitats
- Shelf habitats
- Oceanic / beyond shelf habitats

Variable salinity habitats were defined according to EU GES Decision 2017/848 for situations where estuarine plumes extend beyond waters designated as Transitional Waters under Directive 2000/60/EC.

Coastal habitats were defined according to EU GES Decision 2017/848. 'Coastal' shall be understood on the basis of physical, hydrological and ecological parameters and is not limited to coastal water as defined in Article 2(7) of Directive 2000/60/EC (WFD).

Shelf habitats were defined according to a mean salinity threshold >34.5 as a boundary between outer coastal and offshore waters, as defined in OSPAR Agreement 13-08, and as was used for nutrients in the previous Common Indicator Assessment IA 2017.

Oceanic / beyond shelf habitats were defined according to a mean depth threshold of >200 m.

The actual integration methods are quite simple by design and should therefore be easy to understand by a general audience. Indicator assessment results should be integrated across common indicators for each of the four pelagic habitat types to ultimately facilitate the determination of a single status for each OSPAR region. The status for each pelagic habitat type should be determined by majority status, however, in the case of a tie when only two indicators are considered common and one of the two indicators has not achieved the threshold or has been assessed by expert judgement to be in "not good" status, the status for the habitat type should also be considered "Not good". In cases where a common indicator is "Not assessed" within a particular pelagic habitat type, the status for that habitat should only consider the indicators which are assessed.

The overall biodiversity status for each OSPAR region is also determined by majority status amongst the pelagic habitats it contains, however, in the case of a tie when only two pelagic habitat types were assessed and one of the two habitat types was in "Not good" status, the status for the habitat type should also be "Not good". In cases where the status for a particular pelagic habitat type is "Not assessed", the status for that OSPAR region should only consider the habitats which were assessed.

2.3 Presentation of results

It is important to display integration results in a table format which has been structured to highlight the hierarchical nature of the integration at the level of pelagic habitat types, as well as the broader OSPAR regions which contain them. Tables should be structured with labels for the four pelagic habitat types nested within OSPAR regions on the left-hand side, followed by the biodiversity status results for each indicator for the corresponding pelagic habitat type. The right-hand side of the table should

mirror the left and display the hierarchical integration of biodiversity status results, first at the level of pelagic habitat types, and finally at the level of each assessed OSPAR region (**Table 3**).

The colour scheme displayed in **Table 2** should be used in order to maintain consistency with the pelagic habitats indicator assessments, as well as the other biodiversity thematic assessments (e.g. benthic habitats), with “Not good” status displayed in red, “Unknown” status in grey, “Good” status in green, and “Not assessed” status white.

When an indicator is not currently considered common, but a pilot assessment for that indicator has been conducted, it is still important to display assessment results, while also clearly identifying that the indicator only has candidate status in a particular OSPAR region. In such cases, corresponding cells in the integration table should be hatched diagonally, following the example in **Table 3**.

Table 3. The status for each indicator within each pelagic habitat type for the three assessed OSPAR regions. Status at the level of pelagic habitat types has been derived from the integration of results for each of the three pelagic habitats indicators. Diagonally hatched cells indicate that an indicator has candidate status in the respective region (i.e. PH3 in Regions II, and IV), and therefore the status of that indicator is not considered as part of the determination of overall status at the regional level.

Region	Habitat	PH1/FW5	PH2	PH3	Habitat status	Region status
Greater North Sea (Region II)	Variable salinity	Unknown	Unknown	Not good	Unknown	Not good
	Coastal	Unknown	Detected trend linked to human activity	Not good	Not good	
	Shelf	Detected trend linked to human activity	Detected trend linked to human activity	Unknown	Not good	
	Oceanic	Not assessed	Not assessed	Not assessed	Not assessed	
Celtic Seas (Region III)	Variable salinity	Unknown	Unknown	Not assessed	Unknown	Not good
	Coastal	Detected trend linked to human activity	Detected trend linked to human activity	Detected trend linked to human activity	Not good	
	Shelf	Detected trend linked to human activity	Detected trend linked to human activity	Unknown	Not good	
	Oceanic	Not assessed	Not assessed	Not assessed	Not assessed	

Bay of Biscay and Iberian Coast (Region IV)	Variable salinity	Not assessed	Not assessed	Not assessed	Not assessed	Not good
	Coastal	Unknown	Detected trend	Not good	Not good	
	Shelf	Detected trend linked to human activity	Detected trend linked to human activity	Unknown	Not good	
	Oceanic	Detected trend linked to human activity	Detected trend linked to human activity	Unknown	Not good	

Using indicator assessment results for the QSR 2023 as a worked example, the overall biodiversity status for variable salinity habitats in the Greater North Sea (OSPAR Region II) was determined to be “Unknown”, since this was the status for both PH1/FW5 and PH2 indicators (**Table 3**). The overall status for coastal habitats was “Not good”, since the PH2 indicator detected a trend linked to human activity (not good), while the PH1/FW5 result was “Unknown”. The overall status for shelf habitats was “Not good”, since both PH1/FW5 and PH2 indicators detected a trend linked to human activity (not good). Considering that coastal and shelf habitats had “Not good” status, while only variable salinity habitats had “Unknown” status, the majority status for pelagic habitats in the Greater North Sea (OSPAR Region II) was “Not good”.

In the Celtic Seas (OSPAR Region III), the overall status for variable salinity habitats was “Unknown” since this was the status for both PH1/FW5 and PH2 indicators, while the PH3 indicator was “Not assessed” (**Table 3**). Due to the “trend detected linked to human activity (not good)” status of coastal habitats under all three pelagic habitats indicators, the overall status for coastal habitats was also classified as “Not good”. The “trend detected linked to human activity (not good)” status of shelf habitats under both PH1/FW5 and PH2 indicators (“Unknown” for PH3) resulting in the classification of the overall status of shelf habitats as “Not good”. Considering that coastal and shelf habitats had “Not good” status, while only variable salinity habitats had “Unknown” status, the majority status for pelagic habitats in the Celtic Seas (OSPAR Region III) is “Not good”.

In the Bay of Biscay and Iberian Coast (OSPAR Region IV), no determination could be made for variable salinity habitats since they were “Not assessed” by the PH1/FW5 and PH2 indicators (**Table 3**). Due to the “trend detected linked to human activity (not good)” status of coastal habitats under PH2, while PH1/FW5 was “Unknown”, the overall status for coastal habitats was also classified “Not good”. Both shelf and oceanic / beyond shelf habitats showed a trend detected linked to human activity (not good) for PH1/FW5 and PH2, therefore the overall status for shelf and oceanic / beyond shelf habitats was “Not good”. Considering that coastal, shelf, and oceanic / beyond shelf habitats were all in “Not good” status, the status for pelagic habitats in the Bay of Biscay and Iberian Coast (OSPAR Region IV) was also “Not good”.

The overall biodiversity status for each OSPAR region (**Table 3**) was also determined by majority status amongst the pelagic habitats it contains, however, in the case of a tie when only two pelagic habitat

types were assessed and one of the two habitat types has not achieved the threshold or was assessed by expert judgement to be in “Not good” status, the status for the habitat type was also “Not good”. In cases where the status for a particular pelagic habitat type was “Not assessed”, the status for that OSPAR region only considered the habitats which were assessed.

2.4 Confidence assessment

A confidence assessment is also conducted according to the recommendations in the QSR Guidance. The overall confidence score for each indicator within each OSPAR region should be determined by considering the type, amount, quality, and consistency of evidence (i.e. Robust, Medium, or Limited), as well as the degree of agreement (i.e. High, Medium, or Low) in the results among the indicator assessments. Confidence is not currently assessed for pilot assessments of candidate indicators since the results of pilot assessments are not a component of the integration of biodiversity status.

First, the level of evidence should be determined for each indicator within each OSPAR region where it is assessed as a common indicator, using a semi-quantitative process based on the criteria described in **Table 4**. Each pelagic habitats indicator assessment includes its own internal confidence assessment, which considers the spatial distribution of samples within each assessment unit, and the temporal distribution of samples within each time-series. Indicator assessments also evaluate the spatial representativeness of the indicator results for each pelagic habitat type. Spatial representativeness is essentially the proportion of assessment units where there were suitable data available to evaluate the indicator assessment, out of the total number of assessment units representing a particular pelagic habitat type. Both the mean confidence in the indicator assessment, as well as the spatial representativeness are used to determine the level of evidence based on expert judgement.

Table 4. Evidence categories and description based on expert judgment. Colours are used for interpretation.

Type, amount, quality, and consistency of evidence	Description
Robust	<p>There are multiple, consistent, independent lines of high-quality evidence</p> <ul style="list-style-type: none"> there are multiple lines of evidence (indicator assessments, other assessments or third-party assessments) with appropriate spatial and temporal scale providing suitable evidence
Medium	<p>There is some high-quality evidence, but gaps remain</p> <ul style="list-style-type: none"> there are multiple lines of evidence (indicator assessments, other assessments or third-party assessments), but the spatiotemporal coverage of those is limited <p>OR</p> <ul style="list-style-type: none"> There are few lines of evidence (indicator assessments, other assessments or third-party assessments), but they do have appropriate spatial and temporal scale providing suitable evidence
Limited	<p>Evidence is limited</p> <ul style="list-style-type: none"> there are few lines of evidence (indicator assessments, other assessments or third-party assessments), and the spatiotemporal coverage of those is limited

Subsequently, the level of agreement should be determined for each indicator within each OSPAR region where it is assessed as a common indicator using the descriptions supplied in **Table 5**. The level of agreement should be evaluated using expert judgement and by considering both the consistency in results across pelagic habitat types for the various components of each indicator (e.g. each of the 8 evaluated lifeforms (PH1/FW5), phytoplankton biomass or zooplankton abundance (PH2), phytoplankton or zooplankton biodiversity (PH3)), as well as the consistency in results across indicator components within each pelagic habitat type (e.g. how similar are the results for phytoplankton and zooplankton).

Table 5. Agreement categories and description as defined by the QSR 2023 guidance. Colours are used for interpretation.

Degree of agreement	Description
High	<p>There is good consensus in the results of the assessments</p> <ul style="list-style-type: none"> the lines of evidence (indicator assessments, other assessments, or third-party assessments) all agree
Medium	<p>The results of the assessments are mostly in consensus but with some deviation</p> <ul style="list-style-type: none"> the lines of evidence (indicator assessments, other assessments, or third-party assessments) mostly agree, although a minor proportion show some deviation
Low	<p>The results of the different assessments do not agree</p> <ul style="list-style-type: none"> the lines of evidence (indicator assessments, other assessments, or third-party assessments) present differing results

To ensure a degree of caution in evaluating confidence in indicator results, for both confidence criteria (i.e. evidence and agreement) a single status was determined by selecting the lowest score amongst the assessed indicators. For example, if PH1/FW5 had an agreement score of “High”, and PH2 had an agreement score of “Low” within the same OSPAR region, then the agreement score for that region would be “Low”. The same approach was followed to integrate an evidence score for each OSPAR region.

Finally, to integrate an overall confidence status for each OSPAR region an overall confidence score is determined by locating the appropriate agreement and evidence scores in the matrix shown in **Table 6**, and selecting a confidence score with the same fill colour in the Overall confidence key below the matrix.

Table 6. A depiction of the evidence and agreement statement and their relationship to the confidence criteria from the QSR Guidance. An overall confidence score is determined by locating the appropriate agreement and evidence scores in the matrix and selecting a confidence score with the same fill colour in the bottom table.

High agreement Limited evidence	High agreement Medium evidence	High agreement Robust evidence
Medium agreement Limited evidence	Medium agreement Medium evidence	Medium agreement Robust evidence
Low agreement Limited evidence	Low agreement Medium evidence	Low agreement Robust evidence

Overall confidence

Very low	Low	Medium	High	Very high
----------	-----	--------	------	-----------

It may be useful to refer to the worked example of confidence assessment from the Pelagic Habitats Thematic Assessment in the QSR 2023 (**Table 7**). In the worked example, agreement and evidence scores were consistent across indicators. By nature of the fact that most indicator assessments share the same source data, it is expected that there will be high consistency in evidence, and likely also agreement, scores across pelagic habitats indicators within each OSPAR region.

Table 7. Confidence assessment of the type, amount, quality, and consistency of evidence (i.e. Robust, Medium, or Limited), as well as the degree of agreement in the results (i.e. High, Medium, or Low) for the three pelagic habitats indicators across the five OSPAR regions. Cells are coloured according to the key in **Table 6**. Diagonally hatched cells indicate that an indicator has candidate status in the region and while a pilot assessment has been conducted, a confidence assessment has not been produced.

Region	Criteria	PH1/FW5	PH2	PH3	Criteria status	Region status
Greater North Sea (Region II)	Agreement	Medium	Medium	Not assessed	Medium	Medium
	Evidence	Medium	Medium	Not assessed	Medium	
Celtic Seas (Region III)	Agreement	High	High	High	High	High
	Evidence	Medium	Medium	Medium	Medium	
Bay of Biscay and Iberian Coast (Region IV)	Agreement	High	High	Not assessed	High	Medium
	Evidence	Limited	Limited	Not assessed	Limited	

3 Change Management

The development and improvement of integration methods is an ongoing process. These integration methods for pelagic habitats thematic assessments have been introduced for the first time to support the QSR 2023. Looking forward to future assessments, it will be increasingly important to adapt a more consistent integration framework across the multiple OSPAR biodiversity assessments in order to aid

interpretation of assessment results. Future work to improve these integration methods will be resource dependent.

Responsibility for follow up of assessment (e.g. if the monitoring is not adequate) and progressing development of the integration method:

- Pelagic Habitats subgroup of ICG-COBAM
- ICG-COBAM
- BDC

4 References

Dierschke, V., A. Kreutle, N. Häubner, C. Magliozzi, S. Bennecke, L. Bergström, A. Borja, S. Boschetti, A. Cheilari, and D. Connor. 2021. Integration methods for Marine Strategy Framework Directive's biodiversity assessments. EUR.

Enserink, L., A. Blauw, D. van der Zande, and S. Markager. 2019. Summary report of the EU project 'Joint monitoring programme of the eutrophication of the North Sea with satellite data' (Ref: DG ENV/MSFD Second Cycle/2016), p. 21 pp.

McQuatters-Gollop, A., L. Guérin, N. L. Arroyo, A. Aubert, L. F. Artigas, J. Bedford, E. Corcoran, V. Dierschke, S. A. M. Elliott, S. C. V. Geelhoed, A. Gilles, J. M. González-Irusta, J. Haelters, M. Johansen, F. C. P. Lynam, N. Niquil, B. Meakins, I. Mitchell, B. Padegimas, R. Pesch, I. Preciado, I. Rombouts, G. Safi, P. Schmitt, U. Schückel, A. Serrano, P. Stebbing, A, and C. Vina-Herbon. 2022. Assessing the state of marine biodiversity in the Northeast Atlantic. *Ecological Indicators* **141**: 109148.