D1.4a: Integration within and across Pelagic Habitats indicators

Matthew Holland[1], Arnaud Louchart[2], Luis Felipe Artigas[2] & Abigail McQuatters-Gollop[1]

[1] Marine Conservation Research Group, University of Plymouth, Drake Circus, Plymouth, PL4 8AA, United Kingdom

[2] Laboratoire d'Océanologie et Geosciences, UMR 8187 LOG, Centre National de la Recherche Scientifique, Université du Littoral Côte d'Opale, Université de Lille, IRD, Wimereux, France

Acknowledgments:

This report was produced as a result of the European Maritime and Fisheries Fund project NEA-PANACEA (North East Atlantic project on biodiversity and eutrophication assessment integration and creation of effective project). The project was funded by the European Union (EU). Grant No. 110661/2020/839628/SUB/ENV.C.2

Disclaimer

List of abbreviations

BDC: OSPAR Biodiversity Committee

EQR: Ecological Quality Ratio

EcApRHA project: Applying an Ecosystem Approach to (sub) Regional Habitat Assessment – an EMFF-funded project

GES: Good Environmental Status

MSFD: Marine Strategy Framework Directive

OO-AO: One Out – All Out

PH: Pelagic Habitats

UKMS: United Kingdom Marine Strategy

Table of contents

1. Background

The goal of the Marine Strategy Framework Directive (MSFD; 2008/56/EC) is to achieve Good Environmental Status (GES) for all European marine waters by 2020, through the evaluation of 11 key descriptors. Currently, the biodiversity assessment of GES in the North-East Atlantic is carried out using a suite of biodiversity indicators; however, robust methods for integrating these indicators to determine biodiversity status have yet to be developed. This ecosystem approach can be supported through synthetising results from multiple related indicators into ecologically relevant indices which are more easily interpreted by policy makers and the public. There is a need to integrate within and across ecosystem components, since confidently establishing quality status relies on having results from a large variety of indicators to assess MSFD criteria for the various descriptors.

The ecosystem approach is fundamental to the MSFD because it provides a holistic view of environmental status. Up until the OSPAR Quality Status Report 2023, the assessment approach for MSFD Descriptor 1 Criteria 6 (D1C6; European Commission, 2017) has been focused on interpreting separate results from the individual biological quality elements that define pelagic habitats through individual pelagic habitats indicator assessments (i.e. changes in phytoplankton and zooplankton communities, in phytoplankton biomass and zooplankton abundance, and in plankton diversity) and applying these separate results to improve implementation of the OSPAR convention (McQuatters-Gollop et al., 2022). However, integrating the results of the Pelagic Habitats indicators hierarchically can also be useful for summarising complex results from multiple indicators for policy makers and the public and for distilling the GES implications of multiple changes occurring within the same marine ecosystem component (i.e. pelagic habitats). The overarching aim of this report is to demonstrate how the Changes in Phytoplankton and Zooplankton Communities, Changes in Phytoplankton Biomass and Zooplankton Abundance, and Changes in Plankton Diversity indicators (PH1/FW5, PH2 and PH3, respectively) from MSFD Descriptors 1 and 4 can be linked through a set of integration rules which evaluate the cumulative effects of multiple indicator results for particular habitat types and ultimately for entire OSPAR regions.

Further, there are clearly strong connections between the PH1/FW5 and PH2 indicators which assess the abundance of plankton lifeforms, and the abundance and biomass of copepods and chlorophyll-*a*, respectively. It is important to understand how these indicators co-vary in order to improve the interpretation of pelagic habitats assessment results.

This report is intended to present and critically evaluate options for a robust integration of multiple pelagic habitats indicator results. We compare results from the PH1/FW5 and PH2 indicators for two assessment areas in the Greater North Sea, we present two options for integrating the results of the indicators which describe changes in phytoplankton and zooplankton communities (PH1/FW5) (Holland et al., 2023), phytoplankton biomass and zooplankton abundance (Louchart et al., 2023a), and in plankton diversity (Louchart et al., 2023b), and we discuss the suitability of threshold values for pelagic habitats assessment.

This deliverable for the NEA PANACEA project (D1.4a) represents the continuation of the integration steps developed during the European Union Directorate-General for Environment (DG ENV) co-funded EcApRHA project (Budria et al., 2017; Elliott et al., 2017), and is closely interrelated with NEA PANACEA D1.5 (Linking Pelagic Habitats indicators with food web indicators and their connection to other ecosystem components and MSFD descriptors), therefore the backgrounds of both reports reference similar concepts and ideas.

2. Relationships between pelagic habitats indicators

2.1. Description of the Pelagic Habitats indicators

2.1.1. PH1/FW5 Indicator - Changes in Phytoplankton and Zooplankton Communities

PH1/FW5 is an indicator of ecosystem function, which measures changes in the abundance of important plankton functional groups or "lifeforms" and determines whether trends in lifeform abundance are associated with parallel trends in pressures. Significant changes in the abundance of planktonic lifeforms represent important changes in ecosystem functioning and have consequences for food webs and trophic interactions.

For the QSR 2023, eight lifeforms were highlighted, due to their ecological relevance and owing to the high confidence in their classification (McQuatters-Gollop et al., 2019). Long-term lifeform abundance trends are assessed by calculating the nonparametric Kendall rank correlation coefficient for each time-series, which describes how consistently lifeform abundance has increased or decreased over the time-period assessed. This nonparametric test generates a statistic which is derived by comparing each value in a time-series with each of the values preceding it. The sum of pairwise differences produces Kendall's S-statistic. The variance among these differences is used to derive a Z-score with an approximately normal distribution; thus, confidence in this statistic can be assessed with an associated p-value, with $p \leq 0.05$ generally accepted as statistically significant change. The sign of the test statistic (i.e. positive or negative) reveals the direction of the trend, with a positive statistic indicating an increasing trend and a negative statistic indicating a decreasing trend.

Once statistically significant trends in lifeform abundance have been detected, lifeform abundance time-series are assessed against time-series for relevant environmental parameters to determine whether there are any identifiable correlations which can indicate drivers of change. If variation in an environmental pressure is a strong predictor of lifeform abundance, it is logical to deduce that there is some association between the two time-series. To determine whether any changes in lifeform abundance are associated with parallel changes in these parameters, first monthly time-series for both are smoothed to remove seasonal variation by calculating mean values for each time-series with a 12-month moving window. Subsequently, a random forest algorithm is applied to assess the ability of environmental parameter time-series to predict the lifeform abundance time-series. Parameters are assigned a "variable importance" value representing the net decrease in predictive accuracy if the respective parameter is removed from the model. The result is a ranked list of parameters in order of predictive ability. However, this result alone can not reliably be used to detect whether changes in environmental parameters are linked to changes in lifeform abundance, only that some parameters are better predictors than others. In order to gain a more holistic understanding it is necessary to perform a further integration of indicator results.

For this indicator, no thresholds were available. In the purpose of the QSR 2023, the attribution of quality status was done through applying a set of integration rules explored in section **2.2. Integrating within the PH1/FW5 indicator**. Further information can be in the OSPAR PH1/FW5 indicator assessment (Holland et al. (2023); + web link).

2.1.2. PH2 Indicator – Changes in Phytoplankton Biomass and Zooplankton Abundance

PH2 is a state indicator based on identification of phytoplankton biomass and zooplankton abundance trends within plankton time-series. Anomalies represent deviations from the assumed natural variability of a time-series. Thus, the greater the magnitude of the anomaly (in terms of absolute value, since anomalies can be positive or negative), the greater the change. A value of zero indicates no difference from the time-series mean (which must be de-seasonalised). To understand the changes presented (i.e. annual anomalies) and to be most useful for decision makers, annual anomalies are best interpreted with information provided by anomalies on monthly timescales.

Once the data are at monthly timescale, the time-series analysis can be run. The analysis uses an R-script for both discrete-station data and non-station data. The first step of the indicator analysis consists of identifying the mean seasonal cycle (which is called seasonality in this assessment) during the whole study period. Removing the seasonality is required to analyse the variations of each plankton compartment (i.e. phytoplankton biomass or zooplankton abundance) beyond their natural cycle. The second step consists of obtaining anomalies by subtracting this seasonality from the original time series. The method used is the seasonal differentiation by the seasonal deviation method. Finally, the cumulative sum of these anomalies was produced to detect regime shifts in the time-series for the assessment and comparison periods. A Spearman rank correlation test is then implemented to test the trend of the cumulative sum of the anomalies of the assessment and comparison periods. The results of this test can indicate a significant ($p \leq 0.05$) increase in phytoplankton biomass/zooplankton abundance (0 to 1), no changes (=0) or decrease in phytoplankton biomass/zooplankton abundance (-1 to 0). The results of the Spearman rank correlation provided an indication of change. A t-test against the cumulative sum of the anomalies of the comparison period and the assessment period informs whether the trends are significantly different or not.

For this indicator, no thresholds were available. For the QSR 2023, the attribution of quality status was determined through the application of the One Out – All Out (OO-AO) principle after identifying the relative importance of potential pressures on abundance and biomass anomalies via a random forest modelling approach. Further information can be in the OSPAR PH2 indicator assessment (Louchart et al. (2023a); + web link).

2.1.3. PH3 indicator - Changes in Plankton Diversity

PH3 is a complex multi-metric indicator which describes plankton diversity. The method focuses on α-diversity (i.e. the diversity within a site or sample) and β-diversity, which focuses on the rate of change, or turnover, in species composition (Rombouts et al., 2019). For the QSR2023, we used α- and β-diversity as consecutive steps to detect the temporal changes in community composition (through the β-diversity) and then to report the state of the community where changes were seen (through the α-diversity). First, the beta diversity was computed and significant deviation from the overall composition was flagged. More specifically, the Local Contribution to Beta Diversity (LCBD) shows how much each observation in a time-series contributes to β-diversity. As an example, a site with an average species composition would have an LCBD value near 0. Large LCBD values may indicate sampling units (in time) characterised by high conservation value or degraded and species-poor sites in need of restoration (Legendre and De Cáceres, 2013). High values (approaching a maximum value of 1) may also correspond to special ecological conditions, or may result from the disturbance effect of invasive species (i.e. differing from normal conditions in a positive or a negative way). Whenever a significant change in community composition is detected, the alpha diversity is investigated to observe whether changes in richness and/or dominance were responsible for the result. Assessment of richness is processed using the Menhinick index. The dominance of phytoplankton is assessed using the Hulburt index while the dominance of zooplankton is assessed using the Patten index. Further explanations on the choice of the index can be found in Louchart et al. (2023b).

For the OSPAR Quality Status Report 2023, PH3 was evaluated as a common indicator in the Celtic Seas (OSPAR Region III), and as a candidate indicator in the Greater North Sea and the Bay of Biscay and Iberian Coast (OSPAR Regions II and IV, respectively). While the results of the pilot assessments of PH3 for the Greater North Sea (Region II) and the Bay of Biscay and Iberian Coast (Region IV) were included in the thematic assessment, they were not considered in the region-specific integration for OSPAR Regions II and IV, since PH3 is not currently a common indicator in those regions (**Table 1**). For this reason, results from the PH3 indicator are currently only factored into the integration of overall quality status where PH3 is considered a common indicator (i.e. the Celtic Seas (OSPAR Region III)).

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

Region	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

For this indicator, no thresholds were available. For the QSR2023, the attribution of quality status was determined by identifying of the most important potential pressures on the normalised EQR of β-diversity across assessment units and by the OO-AO principle on the normalised EQR-pressures relationship. This indicator can also be derived through a multi-metric approach which uses a set of biodiversity indices to determine a combined result. Further information can be found in the OSPAR PH3 indicator assessment (Louchart et al. (2023b); + web link).

2.2. Integrating within the PH1/FW5 indicator

Prior to the QSR 2023, the separate components of the PH1/FW5 indicator (i.e. the results for each lifeform) had been assessed individually, or as ecologically meaningful “lifeform pairs” for each spatial assessment unit (e.g. McQuatters-Gollop et al., 2019). With recent developments to the PH1/FW5 Coordinated Environmental Monitoring Programme (CEMP) guidelines, facilitated by the NEA PANACEA project, the integration methodology now groups indicator results by pelagic habitat type to produce a single quality status designation for each habitat type based on results for the 8 high confidence plankton lifeforms highlighted in the current assessment (i.e. diatoms, dinoflagellates, holoplankton, meroplankton, small copepods, large copepods, fish larvae/eggs, and gelatinous zooplankton).

All three pelagic habitats indicators are currently analysed to evaluate biological changes occurring across a set of assessment units and fixed-point stations within the OSPAR maritime area. 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).

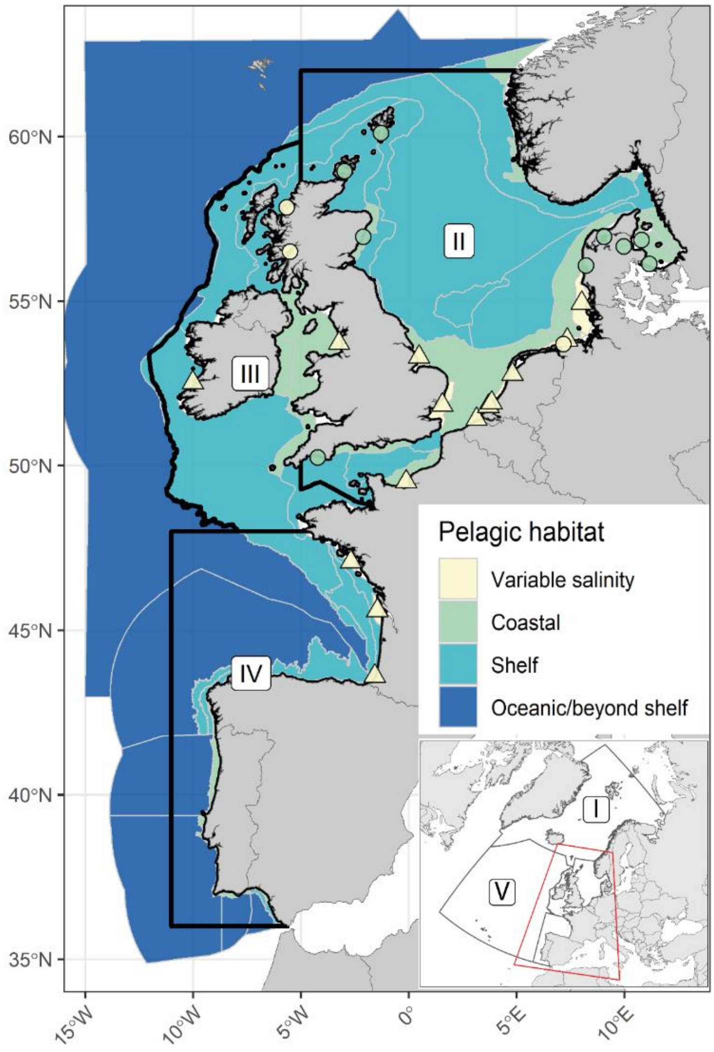

Figure 1. The distribution of the four pelagic habitat types across the three OSPAR Regions considered in the Pelagic Habitats indicator assessments, and the boundaries of the five OSPAR Regions across the OSPAR maritime area (inset). Fixed-point stations are represented as circles and river plumes are represented as triangles. Boundaries between COMP4 assessment units used in this assessment are indicated in grey.

A primary objective of the pelagic habitats assessment for the QSR 2023 was to facilitate an improved understanding of changes occurring coherently across specific pelagic habitat types within 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 at the habitat level (**Figure 1**). The rationale behind this approach was to support the formation of broad overarching conclusions on the status of each pelagic habitat type within each OSPAR region, despite some likely important differences among assessment units of the same habitat type. For a more detailed description of indicator results at the scale of the individual assessment units, consult the individual pelagic habitats indicator assessments (i.e. Holland et al., 2023; Louchart et al., 2023a; Louchart et al., 2023b).

The habitat type categories were created to align assessment outputs with EU MSFD features, with a view of allowing Contracting Parties to use this information for their 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. For future assessment, this salinity threshold will need to be reassessed, as isohaline boundaries will likely be impacted by salinity changes already occurring in the wider Atlantic.

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

There are now several criteria produced by this indicator, which are considered in the current integration methodology. These criteria now include:

- the net trend of Kendall rank correlation coefficients
- a mean confidence score which takes into consideration both spatial and temporal confidence
- how well represented each pelagic habitat type is within the respective OSPAR region
- the most important environmental variable predictor of lifeform abundance

The net trend, or the mode of trends, Kendall rank correlation coefficients describes the primary direction of change detected across assessment units and fixed-point stations within each pelagic habitat category. As an example, changes in meroplankton abundance were assessed across 10 COMP4 assessment units and 2 fixed-point station representing coastal habitats within OSPAR Region II (

Table 2). If there were 0 decreasing trends, 8 increasing trends, and 4 instances of no trend across these assessment units, we would report an increasing net trend and the proportion of assessment units studied where this trend was detected, in this case 8 / 12 = 0.67.

Table 2. Integration of the indicator results for OSPAR Region II – Greater North Sea. Column names are described as follows: ↓: the number of COMP4 assessment units and fixed-point stations where decreasing trends have been detected, -: the number of COMP4 assessment units and fixed-point stations where no trends have been detected, ↑: the number of COMP4 assessment units and fixed-point stations where increasing trends have been detected, Dir: the net direction of change in lifeform abundance (↓: decreasing, ↑: increasing, -: stable), Trend: the percentage of assessment units exhibiting the respective trend, Conf: the mean confidence of datasets considered in the assessment, Change: a logical variable (TRUE/FALSE) to report whether a net trend is likely given the proportion of locations expressing the trend and the confidence and spatial represenativeness scores, Press1: the environmental pressure with the greatest mean rank for the respective trend, Rank1: the mean rank of the environmental pressure indicated under Pres1, nStn: the total number of fixed-point stations considered, totAssess: The total number of COMP4 assessment units and fixed-point stations considered, totCOMP4: The total number of potential COMP4 assessment units for the habitat category, spatialRep: the spatial representativeness score of the analysis. The status of the individual lifeforms are indicated by the colours in the Lifeform column, according to the legend in **Table 3.** The overall status of

the habitat category is indicated by the colour of the first column, which also identifies pelagic habitat types and follows the same colour legend.

Habitat	Lifeform	↓	-	↑	Dir	Trend	Conf	Change	Press1	Rank1	nStn	totAssess	totCOMP4	SpatialRep
Variable salinity	Diatom	0	4	1	-	80%	51%	FALSE	np	3.5	1	5	9	44%
	Dinoflagellate	1	2	2	↑	40%	51%	FALSE	np	2.5	1	5	9	44%
	Holoplankton	1	0	0	↓	100%	30%	FALSE	ph	1.0	0	1	9	11%
	Meroplankton	0	1	0	-	100%	30%	FALSE	psal	1.0	0	1	9	11%
	Large copepods	1	0	0	↓	100%	30%	FALSE	sst	1.0	0	1	9	11%
	Small copepods	1	0	0	↓	100%	30%	FALSE	sst	1.0	0	1	9	11%
	Fish larvae	0	0	1	↑	100%	30%	FALSE	wspd	1.0	0	1	9	11%
	Gelatinous	0	0	0	NA	NA	0%	FALSE	NA	NA	0	0	9	0%
Coastal	Diatom	2	9	8	-	47%	71%	FALSE	psal	3.2	9	19	12	83%
	Dinoflagellate	5	8	6	-	42%	71%	FALSE	ntot	2.3	9	19	12	83%
	Holoplankton	3	8	1	-	67%	59%	FALSE	psal	2.8	2	12	12	83%
	Meroplankton	0	4	8	↑	67%	59%	TRUE	sst	3.1	2	12	12	83%
	Large copepods	3	7	2	-	58%	59%	FALSE	precip	2.1	2	12	12	83%
	Small copepods	2	6	4	-	50%	59%	FALSE	sst	2.2	2	12	12	83%
	Fish larvae	0	6	6	↑	50%	59%	TRUE	psal	3.2	2	12	12	83%
	Gelatinous	0	3	1	-	75%	64%	FALSE	psal	1.0	2	4	12	17%
Shelf	Diatom	1	5	5	↑	45%	74%	FALSE	phos	3.2	0	11	11	100%
	Dinoflagellate	7	3	1	↓	64%	74%	TRUE	wspd	3.9	0	11	11	100%
	Holoplankton	6	4	1	↓	55%	74%	TRUE	sst	2.0	0	11	11	100%
	Meroplankton	0	2	9	↑	82%	74%	TRUE	sst	2.2	0	11	11	100%

Habitat	Lifeform	↓	-	↑	Dir	Trend	Conf	Change	Press1	Rank1	nStn	totAssess	totCOMP4	SpatialRep
🟥	Large copepods	3	8	0	-	73%	74%	FALSE	sst	2.3	0	11	11	100%
	Small copepods	4	4	3	↓	36%	74%	FALSE	sst	1.3	0	11	11	100%
	Fish larvae	1	4	6	↑	55%	74%	TRUE	attn	2.7	0	11	11	100%
	Gelatinous	0	0	0	NA	NA	0%	FALSE	NA	NA	0	0	11	0%

The overall confidence score for this result was calculated as the mean confidence score amongst all assessment units considered for a particular combination of lifeform and pelagic habitat type. We considered COMP4 assessment units and fixed-point stations as equivalent for this part of the integration. For more detailed information on how the confidence score for each time-series is calculated, consult the PH1/FW5 indicator assessment (Holland et al. (2023); + web link).

To report the spatial representativeness of the result, we calculate the proportion of the total number of COMP4 assessment units with indicator results (e.g. 10 in the previous example), out of the total number of possible COMP4 assessment units representing variable salinity habitats within the OSPAR Region, in this case 12 assessment units. Therefore, the spatial representativeness of the result would be 10 / 12 = 0.83. Note that fixed-point station datasets do not contribute to this score.

Finally, to report links to environmental pressures which can drive changes in lifeform abundance contributing to the net trend, we ranked environmental variables for each location based on their relative variable importance, with a value of 1 assigned to the variable with highest importance, 2 to the variable with second highest importance and so on. For assessment units where the Kendall rank correlation coefficient had the same sign as the net trend (i.e. decrease, stable, or increase), we calculated the mean rank of each environmental variable and reported the variable with rank closest to 1.

To ultimately assign a designation of quality status to the individual lifeforms for each pelagic habitat type based on these four criteria we applied a semi-quantitative methodology described in McQuatters-Gollop et al. (2022), developed from the lessons gained 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 3**). Following the criteria outlined in this study, if a pelagic habitat has been assessed, it 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, since 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 eutrophication. Following this logic, the status of 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 is “Not good”.

Table 3. Biodiversity status categories and colours used for the interpretation, by expert judgement, of indicator biodiversity state (McQuatters-Gollop et al., 2022).

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
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.

We considered quality status at the level of each of the 8 high confidence lifeforms within each pelagic habitat within OSPAR Regions and integrated the results for multiple lifeforms to assign a single quality status designation for the pelagic habitat type. For the status of a lifeform to be shifted from “Unknown” to “Not good” the results of the integration had to meet certain criteria:

- The net trend must either be increasing or decreasing and must be present in at least 50% of the locations assessed.
- The mean confidence for locations considered for the determination of the net trend must be at least 50%.
- The spatial representativeness of the assessed locations out of the total number of possible locations for that habitat type must be at least 50%.
- The environmental pressure with the greatest mean rank for locations expressing the net trend must be linked to anthropogenic pressures (i.e. either sea surface temperature, pH, or nutrients).
- The mean rank of the most important environmental pressure must be ≤ 3, indicating that across all assessment units the variable ranks in the top 3 most important for predicting the abundance of the lifeform.

If all the above criteria are met, the lifeform is assigned a status of “Not good”. If 25% or more of assessed lifeforms within a pelagic habitat type are assigned a status of “Not good” then the whole habitat type is also assigned a status of “Not good”. Otherwise, the habitat type is assigned a status of “Unknown”.

2.3. Integrating across pelagic habitats indicators

2.3.1. Comparison of PH1/FW5 and PH2 indicators

Results from the PH1/FW5 and PH2 indicators were compared within the Coastal Well Mixed (CWM) and Northern North Sea (NNS) COMP4 assessment units, two areas with thorough coverage from the CPR survey (**Figure 2**). CPR data were used to calculate monthly abundance values for the small copepod and large copepod lifeforms (PH1/FW5) so that they could be compared to total copepod abundance (PH2). These results were all generated from the same CPR dataset, so a high degree of correlation should be expected. Similarly, the abundance of diatom and dinoflagellate lifeforms (PH1/FW5) was compared to chlorophyll-*a* biomass (PH2), a proxy for phytoplankton biomass. In this case, chlorophyll-*a* biomass was calculated from remotely sensed satellite data which were available from 1998 to present. To ensure comparability between the phytoplankton and zooplankton analyses, the zooplankton time-series were also limited to a start date of 1998.

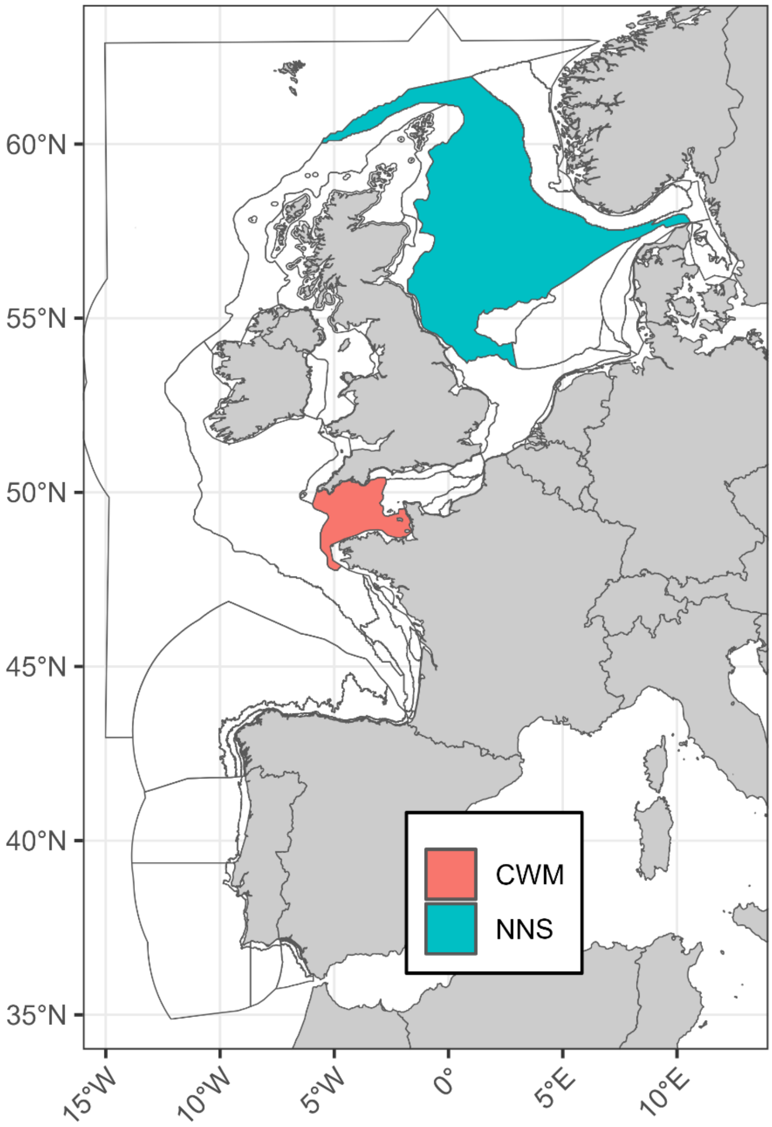

Figure 2. COMP4 assessment units which were used as a case study for the comparison of PH1/FW5 and PH2 indicator results. The COMP4 assessment units selected were Coastal Well Mixed (CWM) and Northern North Sea (NNS).

PH1/FW5 and PH2 monthly indicator results were normalised by subtracting the mean value from each time-series and PH2 results were scaled so they could be displayed on the same y-axis (**Figure 3**). The time-series for diatoms and dinoflagellates in both assessment units were plotted alongside time-series for chlorophyll-*a* in order to study the synchronisation between phytoplankton indicators. In both CWM and NNS the time-series for diatoms showed strong synchronisation with the chlorophyll-*a* time-series up until about 2010. Further, the time-series show very little lag between them, suggesting that the seasonal dynamics of chlorophyll-*a* concentration are being driven by diatoms, as would be expected. Dinoflagellate abundance showed much less synchronisation with chlorophyll-*a*, and their annual bloom periods appear to lag slightly behind the chlorophyll-*a* peak. There were three years of spikes in chlorophyll-*a* which were not reflected in either the diatom or dinoflagellate time-series.

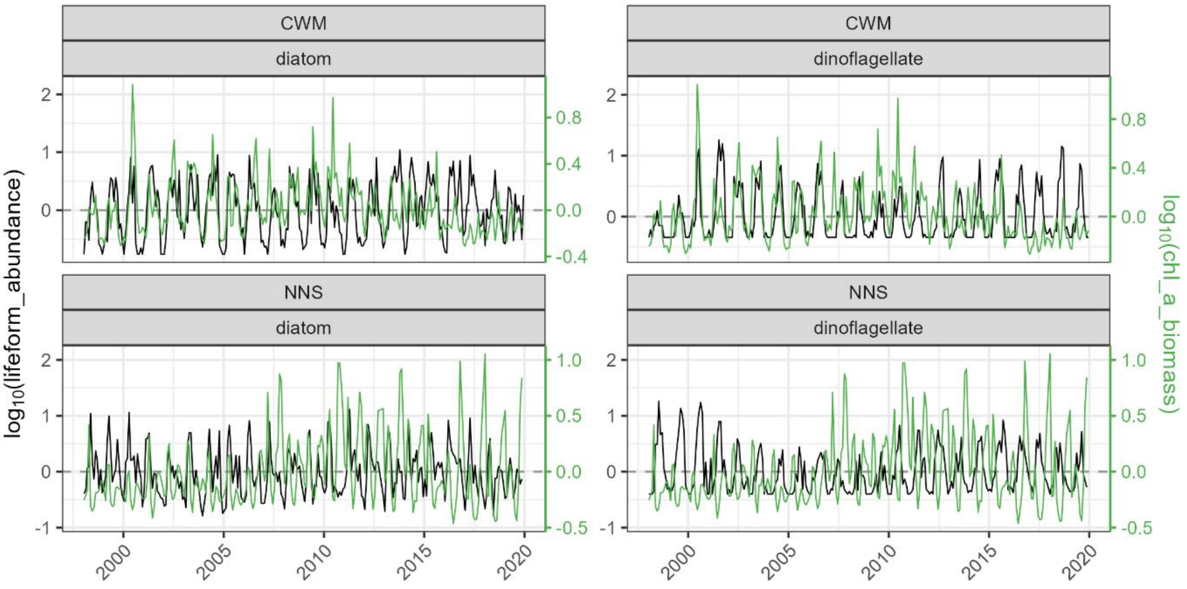

Figure 3. Normalised time-series for diatoms and dinoflagellate (black) in the Channel Well Mixed (CWM) and Northern North Sea (NNS) COMP4 assessment units, overlayed with the time-series for chlorophyll-*a*. The dotted horizontal line indicates the mean value for each time-series. Note that chlorophyll-a has been plotted with a different y-axis scale.

The time-series for large copepods and small copepods were plotted alongside time-series for total copepod abundance to study the synchronisation between zooplankton indicators (**Figure 4**). In both CWM and NNS the time-series for large copepods and small copepods were highly synchronised with the total copepods time-series. This is not unexpected, since the indicators were calculated from the same CPR data.

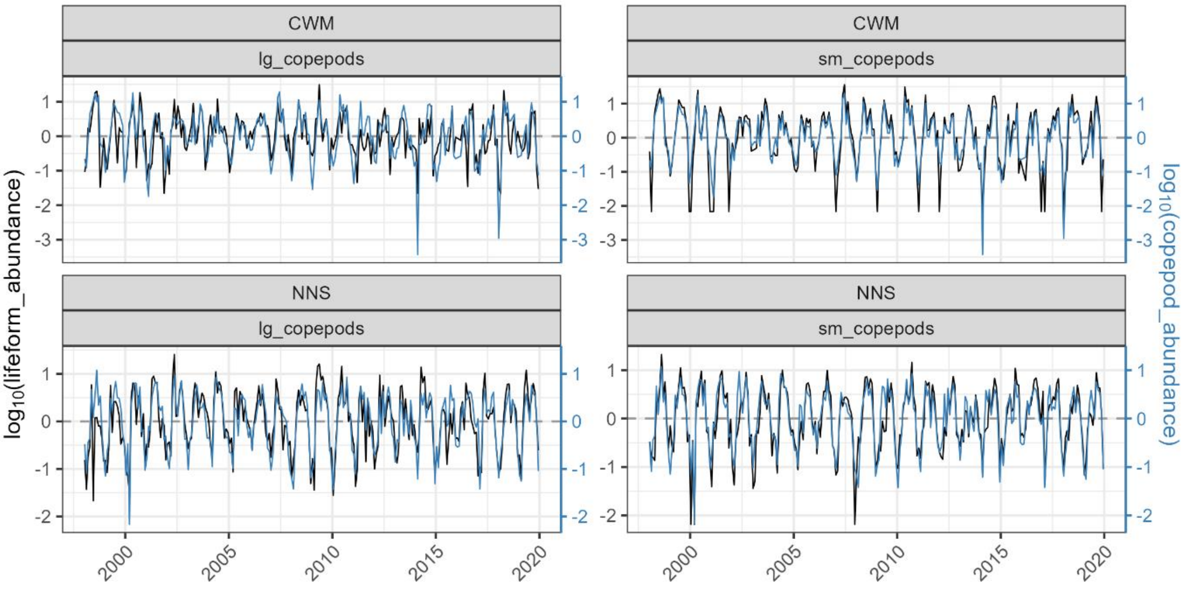

Figure 4. Normalised time-series for large copepods and small copepods (black) in the Channel Well Mixed (CWM) and Northern North Sea (NNS) COMP4 assessment units, overlayed with the time-series for total copepod abundance. The dotted horizontal line indicates the mean value for each time-series.

PH1/FW5 and PH2 monthly indicator results were normalised by subtracting the mean value from each time-series and dividing by the standard deviation so that the annual cycle in each indicator could be compared directly. The annual cycle in chlorophyll-*a* biomass in CWM closely followed the cycle for diatom abundance, with the spring bloom evident (**Figure 5**). The late summer peak in dinoflagellate abundance was also reflected in the chlorophyll-*a* time-series. In NNS, chlorophyll-*a* more closely followed the cycle in diatoms, and the late summer peak in dinoflagellate abundance was poorly reflected.

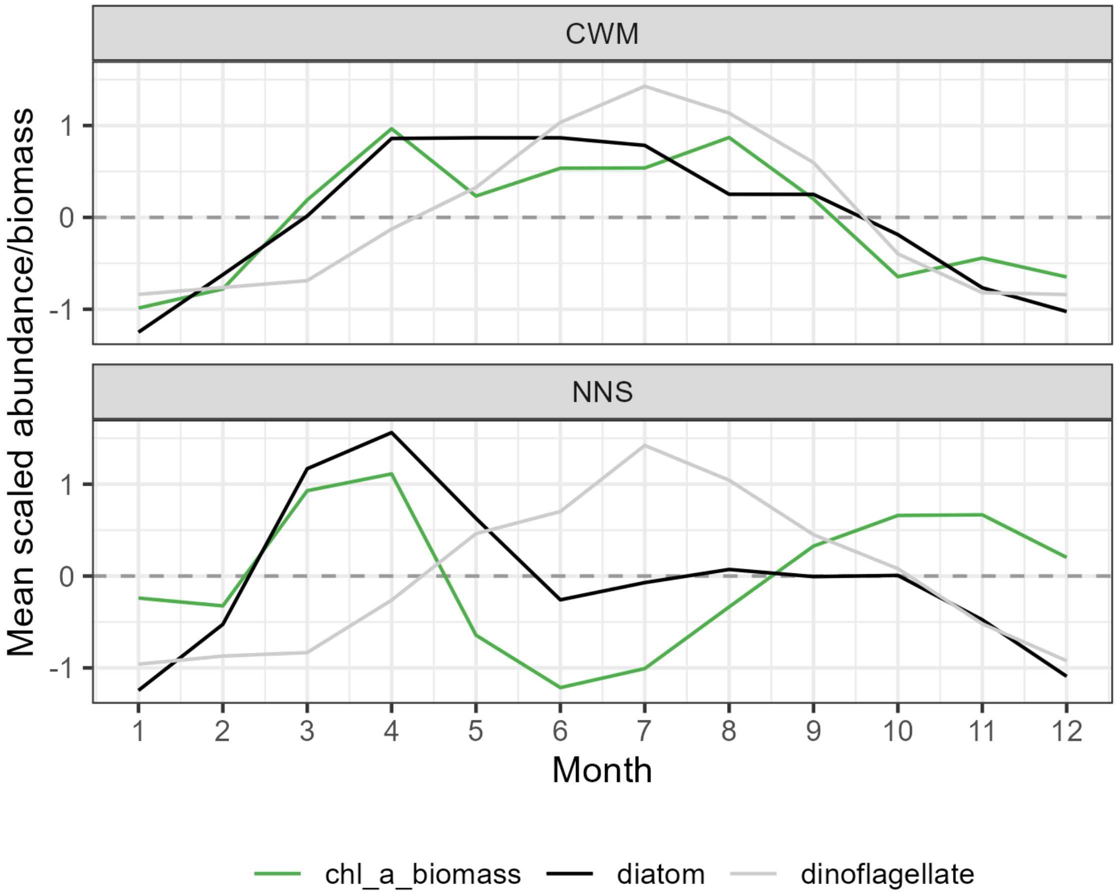

Figure 5. Mean scaled annual cycle for diatom and dinoflagellate abundance and chlorophyll-*a* biomass in the Channel Well Mixed (CWM) and Northern North Sea (NNS) COMP4 assessment units.

The annual cycle in total copepod abundance more closely followed the annual cycle in small copepod abundance, rather than large copepod abundance, particularly in CWM (**Figure 6**). In NNS, the total abundance of copepods remained relatively stable throughout the growing period as a result of the spring peak in large copepods and late summer peak in small copepods.

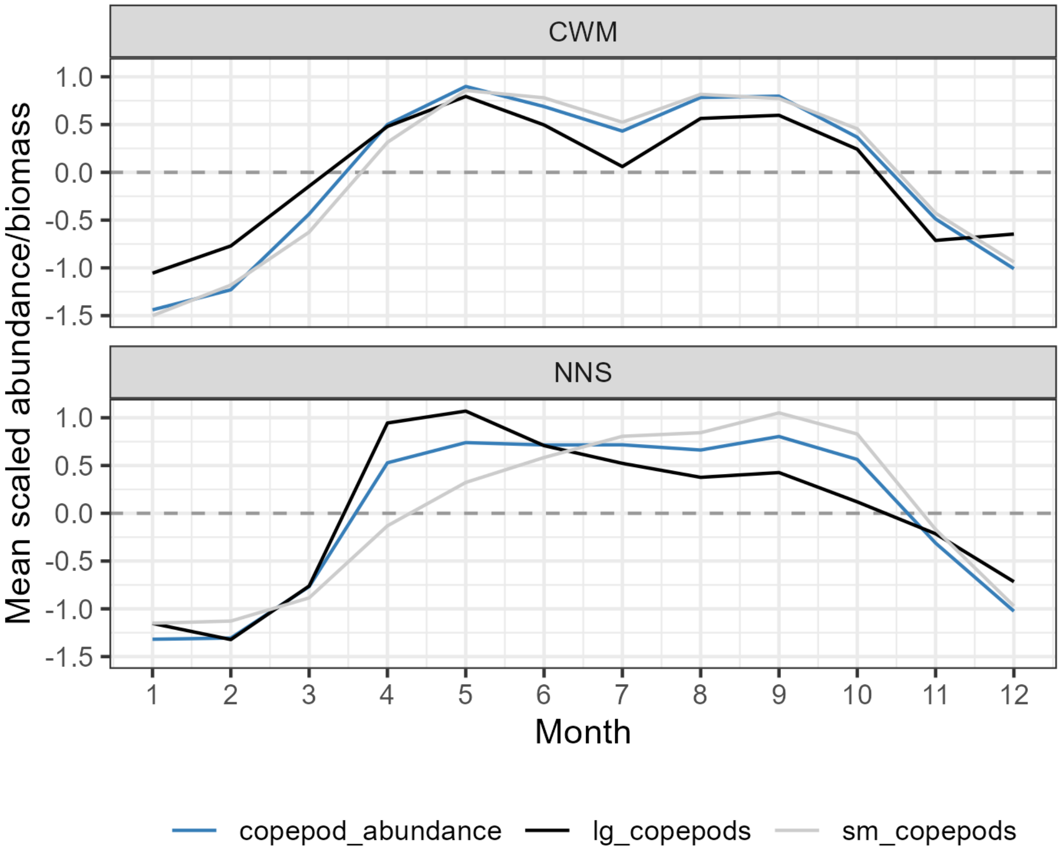

Figure 6. Mean scaled annual cycle for large copepod and small copepod abundance and total copepod abundance in the Channel Well Mixed (CWM) and Northern North Sea (NNS) COMP4 assessment units.

Monthly abundances of phytoplankton lifeforms from the CPR survey were regressed against chlorophyll-*a* biomass from satellite using second order polynomial equations, since maximum primary productivity is typically obtained at intermediate levels of phytoplankton abundance. In all four cases the regression was statistically significant (**Figure 7**). However, in NNS dinoflagellate abundance was negatively correlated with chlorophyll-*a* biomass. While positive correlation was detected in three cases, R^2 values were very low, ranging from 0.05 to 0.17.

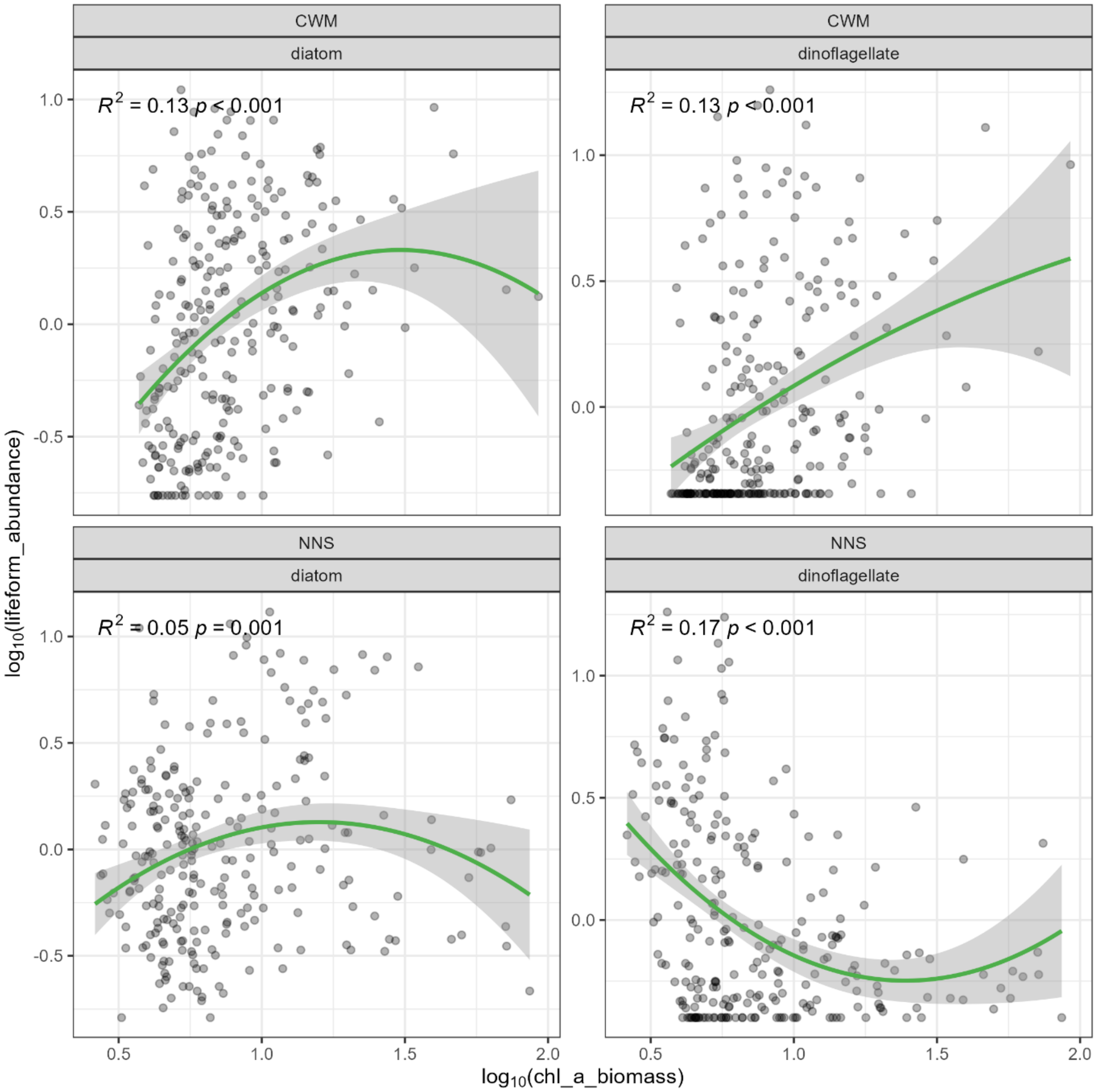

Figure 7. Second order polynomial correlations between the abundance of phytoplankton lifeforms and chlorophyll-a biomass. Trend lines indicate a linear model relationship between variables. R^2 and p-values from Pearson correlation tests are indicated in each plot.

Monthly abundances of zooplankton lifeforms from the CPR survey were regressed against total copepod biomass from the same dataset using linear regression. Significant positive correlation was found in all cases, as would be expected for correlations generated from the same data (**Figure 8**). Unlike for phytoplankton, R^2 values for zooplankton correlations were all very high, ranging from 0.57 to 0.82. The abundance of small copepods showed stronger correlation with total copepod abundance (larger R^2 values) than with large copepod abundance, likely due to the fact that small copepods contribute a greater proportion of total copepod abundance.

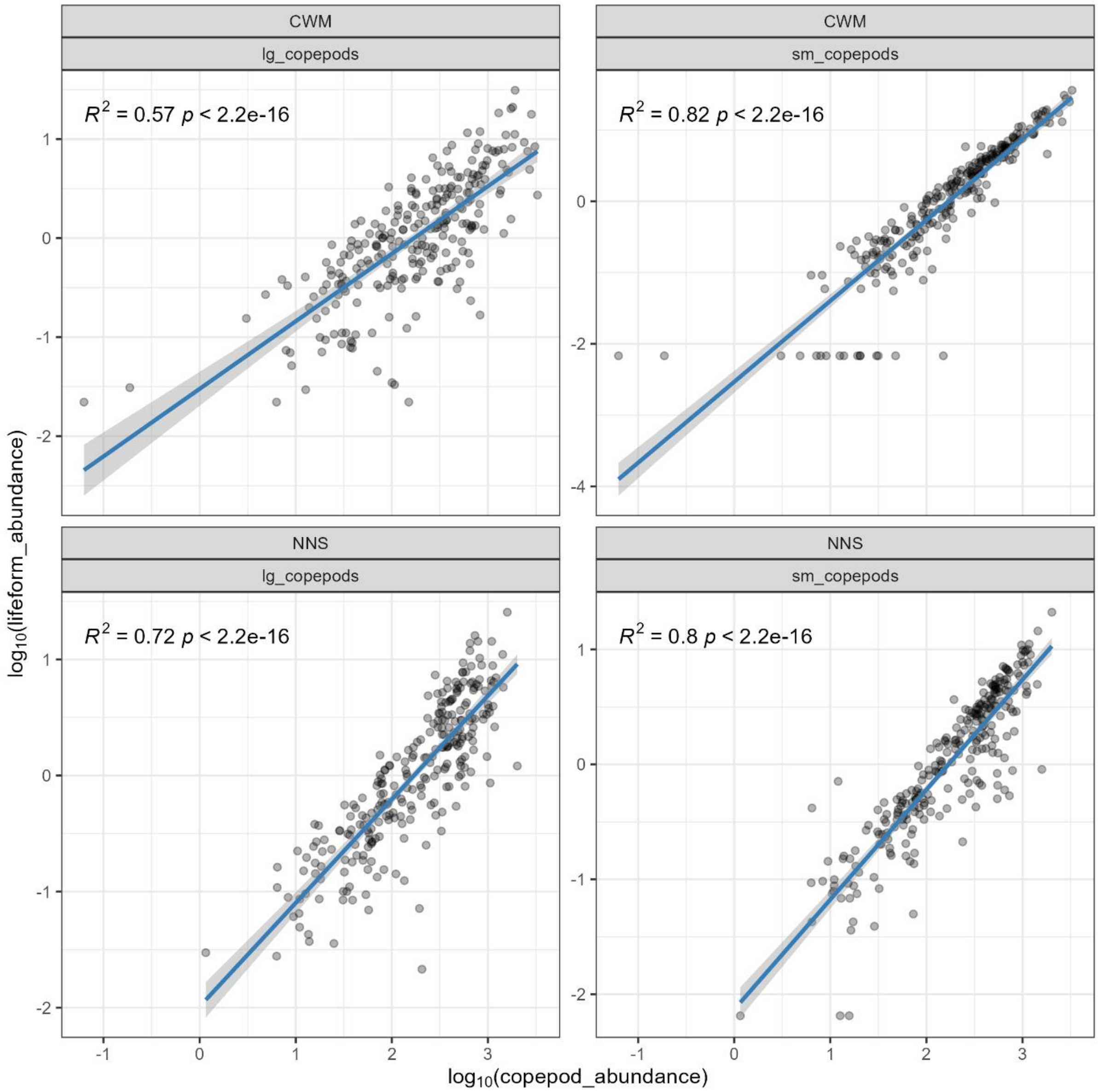

Figure 8. Linear correlation between the abundance of zooplankton lifeforms and total copepod abundance. Trend lines indicate a linear model relationship between variables. R^2 and p-values from Pearson correlation tests are indicated in each plot.

The lifeform pairs indicator, which assesses differences in the relative abundance of two ecologically linked plankton lifeforms, was overlaid with results from the PH2 indicator so that annual cycles in diatom and dinoflagellate abundance could be compared to chlorophyll-a biomass, and large copepod and small copepod abundance could be compared to total copepod abundance. Chlorophyll-*a* biomass in CWM was typically greatest in summer, during periods of high abundance for both diatoms and dinoflagellates (**Figure 9**). In NNS, chlorophyll-*a* biomass was greatest during spring, during periods when typically, diatoms are known to bloom, and dinoflagellate abundance remains low.

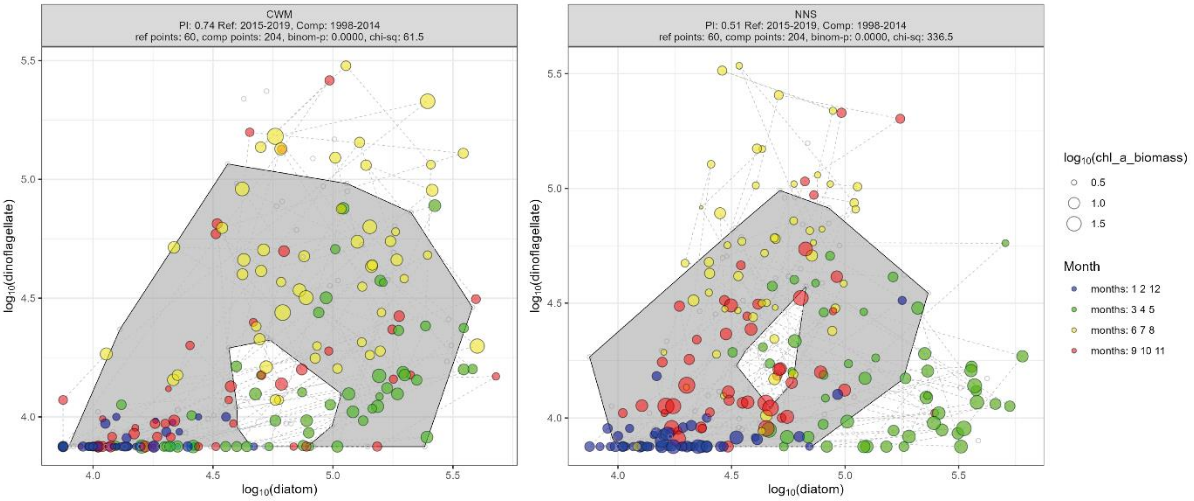

Figure 9. Lifeform pairs indicator output for diatoms and dinoflagellates in the Coastal Well Mixed (CWM) and Northern North Sea (NNS) COMP4 assessment units, overlaid with corresponding results from the chlorophyll-*a* biomass indicator, displayed as point size.

Total copepod abundance was greatest in both CWM and NNS during months when both large copepods and small copepods were at their annual peak in abundance (**Figure 10**). Total copepod abundance was mainly lowest in winter months and highest in summer and autumn.

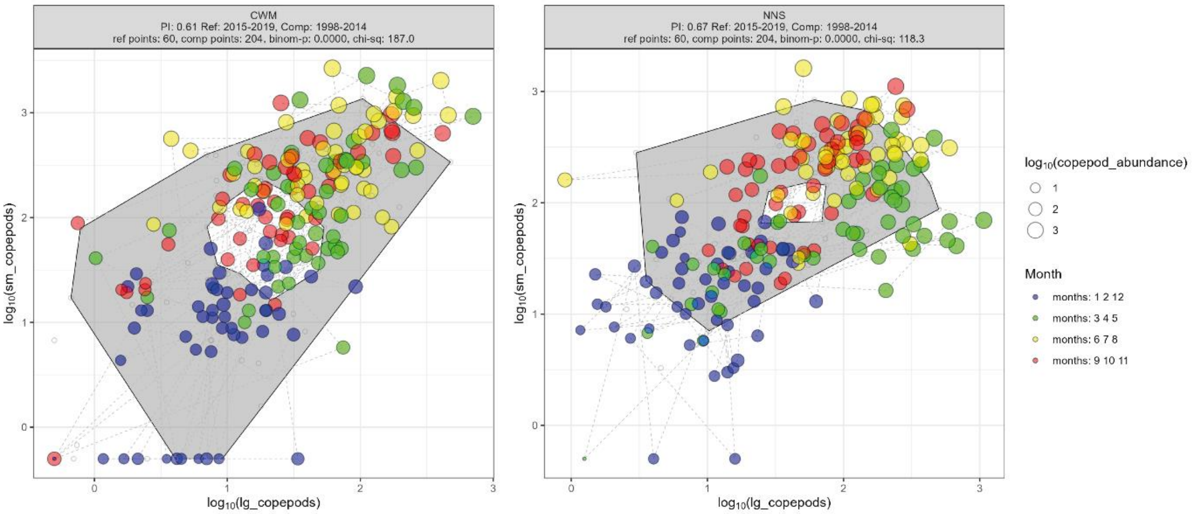

Figure 10. Lifeform pairs indicator output for large copepods and small copepods in the Coastal Well Mixed (CWM) and Northern North Sea (NNS) COMP4 assessment units, overlaid with corresponding results from the chlorophyll-*a* biomass indicator, displayed as point size.

2.3.2. Hierarchical integration options for determining GES

While it can be useful for more technical audiences to interpret indicator results at the level of the four distinct pelagic habitat types that make up each OSPAR region, additional integration may be necessary if a single regional determination of GES is desired. Integration of indicator results is also an important requirement of MSFD and UKMS. The simplest approach would be to extend the rules currently used for integrating results within the PH1/FW5 and PH2 indicators, as described previously (**2.2. Integrating within the PH1/FW5 indicator**). An intermediate step is required to transition from multiple indicator results for each habitat type to a single regional determination of GES for the region; however, there are two possible approaches for this. The first approach would be to integrate the results of the common pelagic habitats indicators for each habitat type to determine an overall quality status result for each habitat (**Figure 11a**). The second approach would be to integrate the indicator results across habitat types to determine an overall quality status result for each indicator (**Figure 11b**). In the next section we test both integration options using results from the OSPAR Quality Status Report 2023 to compare the benefits and drawbacks of these two approaches.

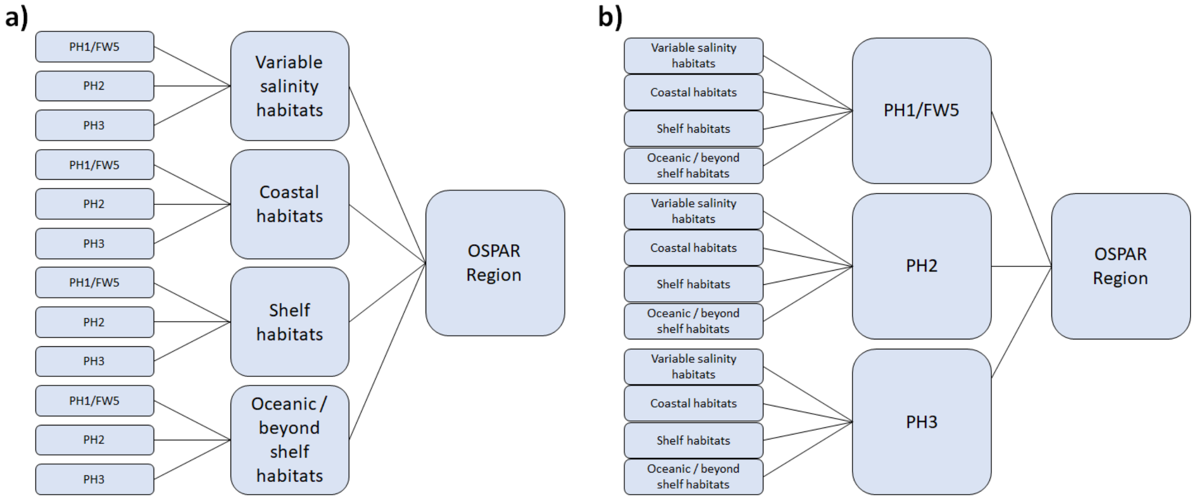

Figure 11. Conceptual flowchart diagram displaying two ways indicator results can be integrated to determine GES. Results can be integrated across the three indicators separately for each of the four four pelagic habitat types within an OSPAR region (a), or they can be integrated across the four pelagic habitat types separately for each of the three pelagic habitats indicators within an OSPAR region (b).

As previously mentioned, it is important to consider that PH1/FW5 and PH2 are closely related indicators and the biological elements forming the basis of PH1/FW5 are the same biological elements which form the basis of PH2. Thus, the PH1/FW5 and PH2 indicators should be provided equal weighting in the integration of GES results.

2.3.2.1. Option 1: Integrating regional biodiversity status by habitat type

The integration of regional biodiversity status by habitat type can be approached using a similar methodology to that which is currently being applied internally to the PH1/FW5 and PH2 indicators, however, certain consideration needs to be given to PH3 indicator results to ensure they only factor into the overall regional status where PH3 is accepted as a common indicator. Similarly, cases when a particular habitat type was not assessed due to not being present within a particular OSPAR region, or due to insufficient data to conduct an assessment, are excluded from the integration.

Across indicator results within the same habitat, majority rules are applied across up to three indicator results (**Table 4**). For Regions II and IV, the maximum number of indicator results to integrate is currently two, however, for Region III the number increases to three due to the inclusion of common indicator results from PH3. In cases where a tie occurs between indicator results for a particular habitat type (e.g. one case of “Unknown” and one case of “Not good”), the overall status for the pelagic habitat defaults to the more negative (i.e. “Not good”, in the previous example) to provide a more cautious assessment result. Although this is somewhat similar to the OO-AO approach, the default determination based on majority status avoids some of its inherent disadvantages (Borja and German Rodriguez, 2010).

To complete the final integration step and achieve a determination of overall status for each OSPAR region, habitat status can be treated in exactly the same way as indicator status, by applying majority rules across pelagic habitat types where the indicator has been assessed.

This methodology was applied to the results of the pelagic habitats assessment for the QSR 2023. When indicator results were integrated by habitat type, variable salinity habitats were designated as "Unknown" in the two OSPAR regions where variable salinity habitats were assessed. All coastal and shelf habitats were designated as having "Not good" status, while oceanic / beyond shelf habitats, only present in Region IV, were also designated as having "Not good" status. The final integration step resulted in all three assessed OSPAR regions being designated as having "Not good" status.

Table 4. The status for each pelagic habitat type within each OSPAR region, derived from integrating the status of common indicators for pelagic habitats. Uncoloured and diagonally hatched cells show that an indicator has candidate status in a particular region and a pilot assessment has been produced. As the PH3 indicator remains a candidate indicator for OSPAR Regions II and IV, the status of PH3 for these regions is given for information purposes only and was not considered in the integration of overall habitat or region status.

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	Not good	Not good	Not good	
	Shelf	Not good	Not good	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	Not good	Not good	Not good	Not good	
	Shelf	Not good	Not good	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	Not good	Not good	Not good	
	Shelf	Not good	Not good	Unknown	Not good	
	Oceanic	Not good	Not good	Unknown	Not good	

2.3.2.2. Option 2: Integrating regional biodiversity status by indicator

The integration of regional biodiversity status by indicator can also be approached using a similar methodology, however, similarly to the previous approach, certain consideration needs to be given to PH3 results. In this case, it is possible to integrate an overall status for PH3 in all OSPAR regions, even where it has only been evaluated as a pilot assessment, since PH3 results are evaluated independently. However, it is important that results for PH3 do not factor into the overall status for the region.

Across the results for the same indicator majority rules can be applied across up to four pelagic habitat types (**Table 5**). In cases where a tie occurs between indicator results for a particular habitat type (e.g. one case of "Unknown" and one case of "Not good"), the overall status for the pelagic habitat defaults to the more negative (i.e. "Not good", in the previous example) to provide a more cautious assessment result.

To complete the final integration step and achieve a determination of overall status for each OSPAR region, habitat status can be treated in the same way as indicator status, by applying majority rules across the overall indicator status result for each region. For Regions II and IV, where PH3 is not considered a common indicator, the PH3 result is ignored during the integration of the overall regional status.

In the below example, this methodology was applied to the results of the pelagic habitats assessment for the QSR 2023. The final integration step resulted in "Not good" status for each indicator, except for in two cases. Similar to the previous example, this integration approach resulted in all three assessed OSPAR regions being designated as having "Not good" status.

Table 5. The status for each pelagic habitat indicator within each OSPAR region, derived from integrating the status of pelagic habitats for common indicators. Uncoloured and diagonally hatched cells show that an indicator has candidate status in a particular region and a pilot assessment has been produced. As the PH3 indicator remains a candidate indicator for OSPAR Regions II and IV, the status of PH3 for these regions is given for information purposes only and was not considered in the integration of overall region status.

Region	Indicator	Variable salinity	Coastal	Shelf	Oceanic	Indicator status	Region status
Greater North Sea (Region II)	PH1/FW5	Unknown	Unknown	Not good	Not assessed	Unknown	Not good
	PH2	Unknown	Not good	Not good	Not assessed	Not good	
	PH3	Not good	Not good	Unknown	Not assessed	Not good	
Celtic Seas (Region III)	PH1/FW5	Unknown	Not good	Not good	Not assessed	Not good	Not good
	PH2	Unknown	Not good	Not good	Not assessed	Not good	
	PH3	Not assessed	Not good	Unknown	Not assessed	Not good	
Bay of Biscay and Iberian Coast (Region IV)	PH1/FW5	Not assessed	Unknown	Not good	Not good	Not good	Not good
	PH2	Not assessed	Not good	Not good	Not good	Not good	
	PH3	Not assessed	Not good	Unknown	Unknown	Unknown	

3. Setting thresholds to facilitate the determination of GES

Although pelagic habitats can in some cases be assessed for GES by evaluating and establishing threshold values based on primary productivity and chlorophyll-*a* (Heyden and Leujak, 2023; Tilstone et al., 2023), there is currently no scientific consensus on what represents GES when it comes to the abundance of lifeforms, copepods, or biodiversity. Further, determination of GES is further complicated by the fact that plankton data collection began after the North-East Atlantic was already heavily impacted by anthropogenic activities, thus there is no data available to represent conditions that are not adversely impacted by human activities.

The current OSPAR methodology used to assess GES for pelagic habitats avoids the use of threshold values. Rather than testing whether a particular threshold value has been attained, the methodology for pelagic habitats indicators developed for the QSR 2023 evaluates three criteria to establish whether there is a suitable burden of evidence, including:

1. a sufficient level of spatial and temporal confidence among assessed time-series,
2. a sufficient level of spatial representation to assess each habitat type, and
3. the most important pressure being one that is linked to anthropogenic activity.

This methodology also evaluates whether the available evidence shows a suitable level of internal agreement to support determination of GES, including:

4. most assessment units showing the same direction of change, and
5. a sufficient mean rank for the most important pressure linked to changes in lifeform abundance).

It could be argued that the minimum levels applied to evidence and agreement criteria are themselves thresholds, however, these values only assess the burden of evidence for whether or not an important change may have occurred, rather than an indicator value above or below which GES is not achieved.

4. Discussion

4.1 Integrating within the PH1/FW5 indicator

Integration within the PH1/FW5 indicator provided a deeper understanding of the larger processes of change occurring in pelagic habitats throughout the North-East Atlantic. The methodology developed for integrating this indicator incorporates important criteria associated with each indicator component and makes a GES decision based on likely links to anthropogenically-influenced pressures. The methodology takes a cautious approach to designating any habitat type as being in “Not good” status by evaluating both the quantity of data available and the level of agreement in results to indicate whether important changes have occurred. This represents a significant advancement in the ability to assess GES for pelagic habitats and avoids the uncertainty and conflicting professional opinions associated with applying threshold values to determine GES.

4.2 Comparison of PH1/FW5 and PH2 indicators

The integration of results from the PH1/FW5 and PH2 indicators provided some additional information beyond that provided by results when interpreted in isolation. Combining these results onto the same plots has exposed seasonal patterns and information about the proportional contributions of each lifeform to bulk phytoplankton and zooplankton.

This process also revealed the poor correlation between lifeform abundance data derived from the CPR survey with chlorophyll-*a* biomass data derived from satellite remote sensing. Although there are several reasons that might explain why remotely sensed values might differ from in-situ measurements, it is likely that differences in the way transect (CPR) versus raster (satellite) data were integrated across assessment units played a part, along with differences in the spatial resolution of the two datasets. Further, our investigation of phytoplankton abundance focused exclusively on micro-phytoplankton. There are important size categories of phytoplankton, including pico- and nano-plankton, which are too small for the CPR to capture effectively which contribute a large proportion of the biomass which is detected by satellite sensors.

It is important to note that relationships between the PH1/FW5 and PH2 indicators have been presented for only two COMP4 assessment units as a case study. Indicator comparisons have not been investigated in depth across the OSPAR maritime area; however, these comparisons have been successful in highlighting important similarities and differences in indicator results. Considering the differences observed between the two assessment units investigated in this report, it is likely that relationships between these two indicators will vary across assessment units, and also across OSPAR regions. The identification of differences and similarities among biodiversity indicators is explored further within NEA PANACEA deliverable D1.5c

Future assessments will likely benefit from adopting an approach which can incorporate the CPR-detected phytoplankton colour index (PCI; a proxy for phytoplankton biomass) in addition to satellite remotely sensed chlorophyll-*a*, to support better alignment of PH1/FW5 and PH2 results. This approach will benefit from inclusion of the longer duration of the PCI time-series (1948 to present), inclusion of the pico- and nano- components of the phytoplankton size spectrum, and co-location of plankton abundance and PCI samples, which will likely facilitate more informative integration results. Further, comparing co-located in-situ measurements of plankton abundance and chlorophyll-*a* biomass, routinely collected from fixed-point monitoring stations, will be useful for identifying how differences in observer platform (i.e. in-situ versus remote sensing) influences indicator results, particularly for coastal regions.

4.3 Hierarchical integration options for determining GES

Between the two options tested to integrate across pelagic habitats indicator results from the current assessment, the overall GES results for the three OSPAR regions were identical. While this identical result is at least partially generated by the high number of indicator results which were in “Not good” status, this is not necessarily a reflection that the integration method is overly sensitive. The results of the indicators themselves had “Not good” status for 12 out of 20 assessed pelagic habitat types (e.g. shelf habitats in Region II).

Although the two options for integration yielded the same GES result for the overall OSPAR regions, it can be argued that one method provides results which are more useful from a management perspective. Integrating across indicators (Option 1) to generate a single GES determination for each habitat type provides granular information on how anthropogenically linked changes vary with geography. These results can inform whether these changes are mainly occurring close to the coast and therefore possibly linked to direct anthropogenic pressures such as eutrophication, or whether they are occurring further offshore and more likely associated with broad scale processes like climate change. By contrast, integrating across habitat types (Option 2) to generate a single GES determination for each indicator produces a result which is not easily interpreted by researchers, policy makers, or the public. Therefore, our recommendation is to use Option 1 for integration, as was used in the QSR 2023.

There are alternate methods for determining GES, which have not been considered in this analysis due to their higher level of complexity. The PH3 indicator uses the EQR approach to determine GES, which integrates multiple indicator components and assigns negative indicator status when particular quantile thresholds are exceeded (Van de Bund and Solimini, 2007). At the present time, the EQR cannot be applied to the PH1/FW5 indicator as it is difficult to define a single indicator value to define a GES threshold. The majority status approach (with OO-AO in the case of a tie) was preferable for this assessment because it promotes transparency through being simple and easily to interpret.

Overall, integrating indicator results for each habitat type presents an easily interpreted framework for summarising the results of multiple indicators. Although the majority status method presents a reasonably simple approach towards integration, it is still important to present the results for each indicator alongside the integrated result wherever feasible to promote transparency. For future assessments we recommend the continued use of the majority status method to determine GES at the level of pelagic habitat types.

5. Conclusion

The integration methods presented in this document represent significant advancements in our ability to assess the status of Pelagic Habitats and generate useful information to compliment the results of each indicator. Simultaneously visualising the results of multiple indicators on the same plot is a valuable tool for understanding how components of the pelagic ecosystem interact. Further, the ability to generate hierarchically integrated GES results provides several nested levels of information suitable for a range of stakeholder audiences, from experts to policy makers, to the public.

At the level of Pelagic Habitats indicators there remains an important knowledge gap about the differential influence of simultaneous pressures, and how variation in pressures can manifest in changes at the community level. To resolve these uncertainties, future assessments will likely need to make use of case studies to evaluate more well-resolved data within assessment units with particularly frequent survey coverage. Finally, there remains a high degree of uncertainty regarding best practices for integrating indicator results in the case of a tie for GES status, since the OO-AO approach tends to over-emphasise negative status (Borja and German Rodriguez, 2010).

6. References

Borja Á, German Rodriguez J. Problems associated with theone-out, all-out'principle, when using multiple ecosystem components in assessing the ecological status of marine waters. PERGAMON-ELSEVIER SCIENCE LTD, 2010.

Budria A, Aubert A, Rombouts I, Ostle C, Atkinson A, Widdicombe C, et al. Cross-linking plankton indicators to better define GES of pelagic habitats. EcApRHA Deliverable 2017.

Elliott S, Arroyo N, Safi G, Ostle C, Guérin L, McQuatters-Gollop A, et al. Proposed approaches for indicator integration. EcApRHA Deliverable WP 4.1. 2017.

Enserink L, Blauw A, van der Zande D, Markager S. 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), 2019, pp. 21 pp.

European Commission. 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 standardised methods for monitoring and assessment, and repealing Decision 2010/477/EU. Official Journal of the European Union L 2017; 125: 43-74.

Heyden B, Leujak W. Winter nutrient concentrations in OSPAR Regions II, III and IV. In: OSPAR, editor. OSPAR, 2023: The 2023 Quality Status Report for the Northeast Atlantic, OSPAR Commission, London, 2023.

Holland M, Louchart A, Artigas LF, Mcquatters-Gollop A. PH1/FW5 - Changes in phytoplankton and zooplankton communities. In: OSPAR, editor. The 2023 Quality Status Report for the Northeast Atlantic, OSPAR Commission, London, 2023.

Legendre P, De Cáceres M. Beta diversity as the variance of community data: dissimilarity coefficients and partitioning. Ecology letters 2013; 16: 951-963.

Louchart A, Artigas LF, Holland M, Mcquatters-Gollop A. PH2 - Change in phytoplankton biomass and zooplankton abundance. In: OSPAR, editor. The 2023 Quality Status Report for the Northeast Atlantic, OSPAR Commission, London, 2023a.

Louchart A, Artigas LF, Holland M, Mcquatters-Gollop A. PH3 - Changes in Plankton Diversity. In: OSPAR, editor. The 2023 Quality Status Report for the Northeast Atlantic, OSPAR Commission, London, 2023b.

McQuatters-Gollop A, Atkinson A, Aubert A, Bedford J, Best M, Bresnan E, et al. Plankton lifeforms as a biodiversity indicator for regional-scale assessment of pelagic habitats for policy. Ecological Indicators 2019; 101: 913-925.

McQuatters-Gollop A, Guérin L, Arroyo NL, Aubert A, Artigas LF, Bedford J, et al. Assessing the state of marine biodiversity in the Northeast Atlantic. Ecological Indicators 2022; 141: 109148.

Tilstone GH, Land PE, Pardo S, Kerimoglu O, Van der Zande D. Threshold indicators of primary production in the north-east Atlantic for assessing environmental disturbances using 21 years of satellite ocean colour. Science of The Total Environment 2023; 854: 158757.

Van de Bund W, Solimini AG. Ecological Quality Ratios for Ecological Quality Assessment in Inland and Marine Waters. RBECCA Deliverable 10 2007.