Report on the integration of OSPAR Food Webs Indicators into the NEAT tool

EcApRHA Deliverable WP 3.5

Co-financed by the European Union

2017

Contents

Authors

Matilda Haraldsson[1]
Nina Larrissa Arroyo[2]
Elisa Capuzzo[3]
Pascal Claquin[1]
Jacco Kromkamp[4]
Nathalie Niquil[1]
Clare Ostle[5]
Izaskun Preciado[2]
Georges Safi[1]

[1]CNRS, UMR 7208 BOREA, Biologie des Organismes et Ecosystèmes Aquatiques, Université de Caen Basse-Normandie, IBFA, Esplanade de la Paix, CS 14032, 14032 Caen Cedex 5, France

[2]IEO, Instituto Español de Oceanografía, Promontorio de San Martín s/n, 39004 Santander, Cantabria, Spain

[3]CEFAS, Lowestoft Laboratory, Pakefield Road, Lowestoft NR33 0HT, UK

[4]NIOZ, Royal Netherlands Institute for Sea Research, PO box 140, 4400 AC Yerseke, The Netherlands

[5]SAHFOS, Sir Alister Hardy Foundation for Ocean Science, The Laboratory, Citadel Hill, Plymouth, PL1 2PB, UK

EcApRHA

The EcApRHA project (Applying an Ecosystem Approach to (sub) Regional Habitat Assessment) aims to address gaps in the development of biodiversity indicators for the OSPAR Regions. In particular, the project aims to overcome challenges in the development of indicators relating to the MSFD (Marine Strategy Framework Directive 56/2008/EU), such as Descriptor D1 (Biodiversity), D4 (Food webs) and D6 (Seafloor integrity), and to deliver an action plan to OSPAR that will enable monitoring and assessment at the (sub) regional scale, to contribute to OSPAR Intermediate Assessment 2017.

Indicators related to the benthic and pelagic habitats, as well as food webs, are investigated within the project at different levels (from data to indicator; from indicator to habitat assessment; from habitat to ecosystem assessment).

Acknowledgment

This report was produced as a result of the EcApRHA (Addressing gaps in biodiversity indicator development for the OSPAR Region from data to ecosystem assessment: Applying an ecosystem approach to (sub) regional habitat assessments) project. The project was co-financed by the European Union (EU). Grant No. 11.0661/2015/712630/SUB/ENVC.2 OSPAR

Institutions and experts that collaborated in this work:
AWI (DE): Ragnhild Asmus; CEFAS (UK): Christopher Lynam; CNRS (FR): Felipe Artigas, Laetitia Petit, Pauline Vouriot, Sandrine Serre; DEVOTES Project experts [AZTI (ES) and Marilim aquatic research (DE)]: Angel Borja and Torsten Berg; MNHN (FR): Isabelle Rombouts; University of Hull (UK): Rodney Forster

Disclaimer

Executive Summary

The main goal within the EU Marine Strategy Framework Directive (MSFD) is to achieve an ecosystem "Good Environmental Status" (GES). To go from indicator to ecosystem assessment, the Food Webs group of the EcApRHA project (applying an Ecosystem Approach to (sub) Regional Habitat Assessments) suggested to use the NEAT tool (http://www.devotes-project.eu/neat/). NEAT is a free software tool designed to run a holistic status assessment including multiple indicators. To demonstrate its feasibility, our focus is on the integration of two of the food webs indicators being developed for use within the North-East Atlantic through OSPAR to support MSFD implementation "primary production" and "mean trophic level". It should be noted that in principle more indicators could be added in future.

To run an assessment using multiple indicators in NEAT, the critical information needed is the definition of a reference value for each indicator, which indicates GES, suggestively through expert knowledge. It is worth noting that the reference values that are proposed in this report are calculated for the purposes of demonstrating the NEAT tool and do not represent any agreement of reference values by European Member States within OSPAR.

The result of the NEAT assessment is a value that indicates the environmental status for each spatial assessment unit, which is a value integrating all indicators considered. Per our present knowledge, the NEAT tool is currently one of the best statistical tools available for integrating multiple indicators. The biggest challenge encountered was to define the requested reference value and surrounding boundaries for each indicator, which involved integrating information from experts' knowledge and literature. It is worth noting however, that within the current revised Commission Decision (2010/477/EU), the European Commission expects each member state to define the reference value for GES (https://ec.europa.eu/info/law/better-regulation/initiatives/ares-2016-5301702_en).

List of Acronyms

BalticBOOST	DG ENV/MSFD co-financed Baltic Sea project to boost regional coherence of marine strategies through improved data flow, assessments, and knowledge base for development of measures
DEVOTES	Development Of innovative Tools for understanding marine
EcApRHA	OSPAR's Ecosystem Approach to (sub) Regional Habitat Assessments project
GES	Good Environmental Status
HELCOM	Baltic Marine Environment Protection Commission - Helsinki Commission
ICES	The International Council for the Exploration of the Sea
ICG COBAM	Intercessional Correspondence Group on the Coordination of Biodiversity and Monitoring
MTL	Mean Trophic Level
MSFD	Marine Strategy Framework Directive
NEAT	Nested Environmental status Assessment Tool
OSPAR	The Convention for the Protection of the Marine Environment of the North-East Atlantic
PP	Primary Production
TL	Trophic level

1 Introduction

Food webs are a simplified representation of species interactions in a specific area or ecosystem. In other words, they are diagrams that depict species trophic interactions in a community, i.e. who eats whom. The proposed full suite of OSPAR food webs indicators (Table A) aims to capture the complexity of food webs by assessing the status of their ecosystem components. Linking information between these indicators supports a holistic description of the food webs. However, in the current state of knowledge, establishing relationships between all food webs indicators is a challenging issue as each of these indicators can be sensitive to a different kind of pressure. Integrative tools are hence needed to allow the aggregation of the various information, in a comparable and systematic way, towards assessing the food webs status. The Nested Environmental status Assessment Tool (NEAT) software implementing a biodiversity assessment is used here.

Table A: List of OSPAR food webs indicators. Three indicators are currently adopted as common indicators (i.e. FW3, FW4 and PH1/FW5).

OSPAR Food Webs code	Indicators
FW1	Reproductive success of marine birds in relation to food availability
FW2	Production of phytoplankton
FW3	Size composition in fish communities (TyL)
FW4	Changes in average trophic level of marine predators (cf MTL)
PH1/FW5	Change of plankton functional types (life form) index Ratio
FW6	Biomass, species composition and spatial distribution of zooplankton
FW7	Biomass and abundance of functional groups
FW8	Changes in the distribution of biomass and species over trophic levels and body size
FW9	Ecological Network Analysis indicator (e.g. trophic efficiency, flow diversity)

GENERAL REMARK FOR THE CONSISTENCY WHILE READING

In the current deliverable, the version 1.2 of the NEAT software was tested to integrate two food webs indicators for the assessment of a sub-regional case study (i.e. Bay of Biscay continental shelf). The main information requested for inclusion in the NEAT tool was the status class boundaries of each indicator. These boundaries include the reference value of the indicator setting the limit between a “Good” and a “Moderate” status of the environment but also additional boundaries between “Good” and “High” status, “Moderate” and “Poor” status and “Poor” and “Bad” status. Several terms refer to the “reference value” such as “reference point”, “reference level” or “assessment threshold”. These values are based on expert calculations for the purposes of this study and do not represent any agreement of reference values within OSPAR. For more consistency in this document, we refer to:

- The assessment value: is the calculated indicator value that needs to be included in the NEAT tool along with its associated standard error. The assessment value will be compared to the reference value to evaluate the state of the ecosystem

- The assessment NEAT value: is the final assessment result of a spatial assessment unit that is given by NEAT tool (i.e. standardized value between 0 and 1)

- The reference value: is the boundary set between “Good” and “Moderate” status

- The two reference boundaries: is related to indicators which need two level of boundaries (i.e. Upper and lower limitations/boundaries for the indicator trend).

- The boundaries: as the different boundaries set between the five status classes in NEAT tool (i.e. “Bad”, “Poor”, Moderate”, “Good”, “High”)

The rest of the terms used in this document are defined in the Glossary part (section 7 of the document).

2 Background

The main goal within the EU Marine Strategy Framework Directive (MSFD) is to achieve a “Good Environmental Status” (GES) of marine waters by 2020 (EU, 2008). To help member states to interpret and achieve GES, the European Commission has proposed the use of criteria and indicators (currently 56 indicators, divided into 11 descriptors) (EU, 2010). As part of the monitoring work of these indicators each member state must determine, based on scientific knowledge, the reference values of each indicator for their own marine waters. The OSPAR Convention is the mechanism by which 15 Governments & the EU cooperate to protect the marine environment of the North-East Atlantic. Currently eleven of the fifteen governments, party to OSPAR are EU Member States and make use of the coordination platform that OSPAR provides to support the regional implementation of relevant EU legislation.

The EcApRHA project (applying an Ecosystem Approach to (sub) Regional Habitat Assessments) was embedded in the work of OSPAR through ICG-COBAM (Intersessional Correspondence Group on the Coordination of Biodiversity Assessment and Monitoring). The EcApRHA consortium has worked on developing biodiversity indicators within the OSPAR and the MSFD context. By focusing on habitat related indicators (benthic and pelagic habitats, and food webs), the consortiums’ ambition has been to overcome identified challenges in the development of indicators relating to MSFD implementation.

One of the objectives within the EcApRHA project has been to define how to progress form from assessing individual indicators to regional ecosystem assessment (including several connected habitats). Nested integrated assessments require assessment tools that ensure clarity, transparency and easy handling of the indicators included. The HELCOM BalticBOOST project has undertaken a review of methods and tools for integrated assessment of biodiversity (HELCOM 2016). Upon the eight tools[1] that were evaluated, only the NEAT tool (Nested Environmental status Assessment Tool) was considered as fully aligned with the MSFD requirements. Thus, the Food Webs group has suggested to test the NEAT tool, developed by the DEVOTES project (http://www.devotes-project.eu/neat/). Furthermore, the OSPAR Food Webs group is associated with the DEVOTES project for the development of some of their indicators. This facilitated the exchange with the DEVOTES experts during the testing process of the NEAT tool. During this process, several issues and difficulties have been addressed by the Food Webs expert group related to how these indicators could be integrated into the NEAT tool.

[1] Eight tools are: HELCOM Biodiversity Assessment Tool (**BEAT**), MARMONI biodiversity assessment tool, DEVOTES biodiversity assessment tool (**NEAT**), Biological Health Index (**BHI**), Index of Biological Integrity (**IBI**), Estuarine Fish Community Index (**EFCI**), French estuarine fish index and US EPA Benthic Index.

Aim

The objective of this work was to search for and communicate potential solutions for how to Integrate information from two of the food webs indicators currently under development within the EcApRHA project, (*i*) primary production and (*ii*) mean trophic level, into the NEAT tool. This involved identifying and addressing questions of both a methodological and ecological nature. By illustrating how two different types of indicators can be integrated into the NEAT tool, our aim has been to share our experience with contracting parties through the ICG-COBAM group and to propose a methodological framework on how to proceed in this work.

This work is contributing to the overall integration and aggregation approach for ecosystem assessment (Work Package 4 of EcApRHA). Integrating the ecological information provided by OSPAR pelagic habitats, benthic habitats and food webs indicators is a big challenge. This report is aiming to propose a possible scenario for integrating the various ecological components addressed, by looking at two very different indicators related to food webs structure (MTL) and function (PP).

The NEAT Tool

This section presents a short summary of the NEAT manual. For detailed information, see http://www.devotes-project.eu/neat/.

The Nested Environmental status Assessment Tool (NEAT) is a software tool for running a holistic status assessment using various indicators for an assessment area. It was developed by the DEVOTES project (http://www.devotes-project.eu) and implemented with biodiversity status assessment in mind. The software is not restricted to only this type of assessment but is universal and can be adapted to other types of assessments. The user-friendly software includes several preconfigured settings (e.g., assessment for four European Regional Seas with assigned indicators, following Teixeira *et al*. 2014). However, the software is organised to be configured for other regions and other indicators.

The first step when running an assessment is to define the Spatial Assessment Units (SAU) of interest. As reflected by the software name, the program is designed so that several SAU units can be included in a nested design. For example the North-East Atlantic Ocean can be divided into regions e.g., Celtic Seas or Greater North Sea. These regions can then be further divided into smaller spatially defined sub-regions (e.g., the Southern, Northern or Central part, etc.). In the next step, one or several habitats can be assigned to each SAU (e.g., pelagic, benthic etc.), which are also nested into a hierarchical structure. Various indicators are then assigned to each habitat (or more precisely, to each SAU/habitat combination), connected through their ecological components (e.g., fish, birds, mammals etc.). Finally, the user has to enter the indicator value for each indicator (typically derived from field assessments), and the associated error for each indicator value is demanded. One important character of the software is the weighting system. This is done to the SAU and habitat units to ensure that one indicator (assigned to a specific habitat and SAU) will not dominate the results.

Given this setup, the results (the NEAT value, i.e., the weighted average of all indicators belonging to a specific group) can be evaluated for each SAU, which is typically the unit of assessment. Within this area, the results of the underlying sub-SAU or sub-units (habitats) can be explored. It is also possible to explore the individual NEAT value for each ecological component. In this way, the user can explore potential sensitive areas or components.

For the program to Integrate all indicators, they need to be on the same scale. This is done by assigning them to different status classes, which are then normalised. The current NEAT version allows for a 5-class status scale (“bad”, “poor”, “moderate”, “good”, “high”)(Fig. 1), which the user can define. Each indicator can have their own class boundaries and indicator scale, and these can vary between SAU and habitat.

Future versions of the software are expected to be more flexible in the number of possible classes to choose.

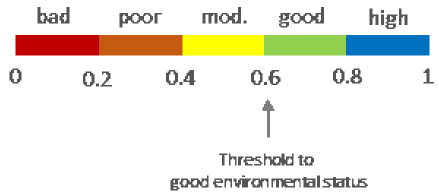


Figure 1: The five-class status scale used in NEAT, with its associated colour code. As a minimum, the lower and higher boundary as well as the reference value (boundary between moderate and good status, which corresponds to good environmental status) must be defined. The figure originates from the NEAT webinar presentation (Berg pers. Com.).

When entering a new indicator (or when modifying an existing one) the user can define these five classes, however, the user is obliged to define at least three of them: the upper and lower boundaries, and the reference value. If only three boundaries are defined, the software will automatically define the in-between boundaries assuming linearity. The GES (Good Environmental Status) is observed from the boundary upwards between “moderate” and “good” status (i.e., the reference value).

To conclude, when running an assessment using multiple indicators in NEAT, the minimum amount of information needed is:

- the indicator value for each assessment unit and habitat (suggestively taken from the field assessment),
- the associated error term to the indicator value,
- the pre-defined minimum boundary of the status classes for each indicator,
- the pre-defined maximum boundary of the status classes for each indicator,
- the pre-defined reference level (boundary between “moderate” and “good” status) for each indicator.

The Indicators

We have mainly focused on two indicators that are currently under development within the Food Web group of the EcApRHA project, in close collaboration with the OSPAR/ICG COBAM Food Web expert group. These include the Mean Trophic Level (MTL) and Primary Production (PP), corresponding to OSPAR indicators FW4 and FW2, respectively.

Mean Trophic Level

The Trophic Level (TL) is the position of an organism in a food web, and the Mean Trophic Level indicator (MTL) reflects the average trophic level of the species present in a food web. MTL is calculated following the equation:

$$MTL_k = \sum_i (TL_i) \times (Y_{ik}) / \sum_i Y_{ik}$$

TLi is the trophic level of species *i* and, *Yik* refers to the biomass of the species in year *k*.

This indicator, which is not specific to fish, is useful to detect effects of fishing pressure on the structure of a food web. As an OSPAR indicator, it is calculated using biomass data from scientific surveys and fisheries landings together with the species assigned TL, derived from dietary analyses (e.g., stomach content, stable isotopes). The group of species considered can be changed by applying a TL cut-off, where the purpose is to eliminate lower TL species (e.g., primary producers and basal resources). Within this report, the indicator that is presented for MTL has been assigned a lower cut-off at TL 3.25. MTL 3.25 focuses on the upper part of the food web since it eliminates all species under a TL of 3.25. It is similar to the Marine Trophic Index (MTI) as proposed by Pauly and Watson (2005), which is currently built into the NEAT tool. The main difference, however, is that the MTL is calculated based on scientific survey data, compared to landing data which are used for the MTI. Hence, the MTL_3.25 includes information directly related to the populations and the unbiased structure of the food web as it is not influenced by fishing variability. Finally, a decrease in MTL 3.25 is typically interpreted as a loss of higher trophic levels, often due to excess fisheries activities. From a sustainability perspective, a decrease in MTL 3.25 is therefore considered as bad, and an increase or a non-significant trend as good.

Phytoplankton Primary Production

Phytoplankton PP is the rate at which phytoplankton produces organic matter. This organic matter will be available for higher trophic levels, and for this reason, PP is fundamental for the structure and function of the ecosystem. In addition, it is highly sensitive and responds quickly to multiple pressures such as nutrient input (Cadee and Hegeman, 2002), grazing, contaminants, non-indigenous species, changes to the physical conditions of the water column (Cole and Cloern, 1987), and climate change. Consequently, the PP indicator is well suited as an early warning signal, and to inform on the general ecosystem status. Using PP as an indicator is challenging due to its two-folded nature. High PP levels are typically associated to strong eutrophication reflecting a probable unhealthy state of an ecosystem. On the contrary, low PP levels are also a sign of external pressures, often due to strong grazing effects or nutrient limitations. Hence, either a significant decrease or increase appears to be characteristic of ecosystem change. Meanwhile, PP levels depend on the region under consideration (e.g., coastal regions are expected to have a higher primary production compared to offshore regions). Therefore, the way to identify these regions is important for future assessment work.

The need to define two reference levels

As explained above, changed trends in PP are important to detect, independent on whether they increase or decrease. Therefore, lower and higher reference boundaries have to be defined. We propose a general status classification (Fig. 2), where boundaries in both directions should be defined, here proposed with the same status-class resolution as in the current version of NEAT (Fig. 1).

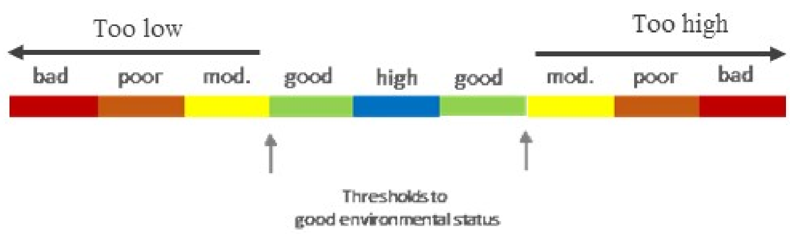


Figure 2: Modifications of the five-class status scale used in NEAT (Fig. 1), including a lower and higher reference boundaries to good environmental status. These classes are proposed as boundary classes for indicators, which have two reference boundaries, i.e., a lower and higher reference value (e.g., Primary Production). Only the upcoming version of NEAT will allow this type of classification (version 1.3 and above). The figure is modified from the NEAT webinar presentation (Berg com pers).

3 Case studies

Several case studies are used to illustrate how scientific data and expert knowledge can be used to define the reference values and boundary levels necessary to integrate an indicator for a specific region into the NEAT tool. These case studies also illustrate some of the potential difficulties encountered in this process. The summaries of the case studies are based on discussions within EcApRHA, including both internal and external experts of the specific cases.

A complete NEAT assessment has been carried out for one SAU (i.e. Bay of Biscay) where MTL and PP data were available and this can be found in annex 1.

Mean trophic level case studies

Trophic Level (TL) estimation for each species can vary in space and time. In order to consider this uncertainty around the MTL indicator, a bootstrap methodology was developed using the R software (R version 3.1.0). A random sampling was applied on TL values and their standard error performing 500 MTL computations per studied year. The model (red line in the Fig. 3 below) was then fitted as a mean value of the 500 MTL generated with an uncertainty related to its standard deviation (Uncertainty is represented by the blue area around the model). The uncertainty around the MTL model is thus linked to the uncertainty of the TL estimation.

Bay of Biscay continental shelf case study (FR)

In this case study MTL3.25 was calculated from survey data (EVHOE survey) from ICES DATRAS Database for the Bay of Biscay continental shelf in France during the period 1997-2014 (ICES region VIIIa and VIIIb). In the region, the overall fishing mortality has been decreasing over the last couple of decades and the biomass of several fish populations have been increasing. The amount of large fish has thus increased (ICES, 2013) and consequently, we observe an increasing MTL 3.25 (Fig. 3).

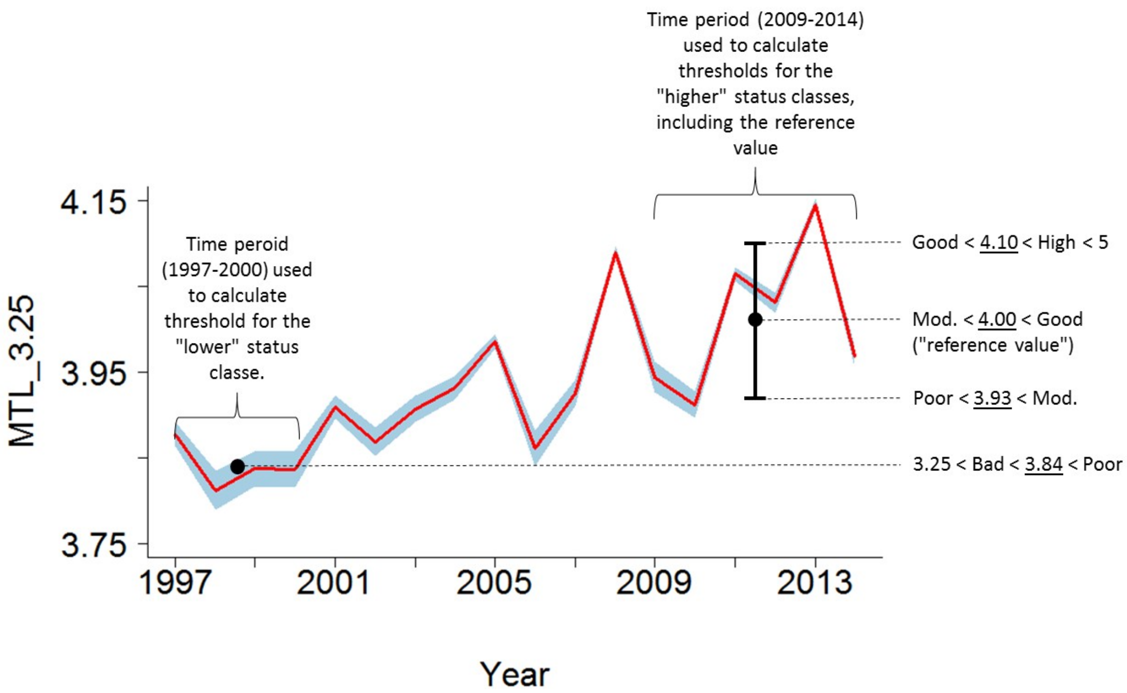


Figure 3: Mean Trophic Level (MTL) with a cut-off of species with trophic level < 3.25 (MTL_3.25) calculated from survey data (red line) from the Bay of Biscay, France, including the standard deviation (blue area) calculated from the uncertainty within species trophic level. The data originates from ICES DATRAS Database (http://www.ices.dk/marine-data/data-portals/Pages/DATRAS.aspx). The figure illustrates the time periods used for calculating the boundary

values for the status classification in NEAT (black dots and error bars). The median value between 1997 and 2000 was used as a boundary between "bad" and "poor" status, and the median value between 2009 and 2014 as the boundary between "moderate" and "good", the later also referred to as the "reference value". The 10th and 90th percentile between 2009 and 2014 was used as the boundary between "poor" and "moderate" (Mod.) and between "good" and "high" status, respectively. The lower value for "bad" status was set to 3.25, and the higher value for "high" status at 5 (table 2).

The reference value for MTL 3.25 was proposed by experts to be calculated from the most recent years that are reflecting a recovering ecosystem in the current case study. This period was chosen following the precautionary principle, since the system may experience further improvement (increase) in the MTL.

As described in previous sections, NEAT requires the user to aggregate the indicator into five status classes ("Bad", "Poor", "Moderate", "Good", and "High" status). For the purpose of this case study, we defined the reference value (boundary between "Moderate" and "Good" status) as the median value between 2009 and 2014 (corresponding to an MSFD-cycle of six years) (Fig. 3). The boundaries surrounding the reference value were based on the 10th and 90th percentile of the same time period (i.e. 2009-2014). The lower boundary value between "Bad" and "Poor" status was based on the median value of the period with the lowest recorded MTL, 1997-2000. This period was chosen since the MTL indicator was at its lowest values reflecting a period where impacts due to fishing were stronger than in the current situation, hence in a less-desirable state (table 1 and 2, Fig. 3). For comparison reasons only, we also calculated the same boundary values based on mean and -/+ standard deviation (table 2). Clearly, the two methods result in very similar boundary levels.

Year	MTL (>3.25)
1997	3.88
1998	3.81
1999	3.84
2000	3.84
2001	3.91
2002	3.87
2003	3.91
2004	3.93
2005	3.99
2006	3.86
2007	3.93
2008	4.09
2009	3.94
2010	3.91
2011	4.06
2012	4.03
2013	4.14
2014	3.97

Table 1: Average yearly Mean Trophic Level (MTL) with a cut-off of species with trophic level < 3.25 (MTL_3.25) from the Bay of Biscay, France. This time series was used to calculate the boundary between the status classes (table 2). Grey sections correspond to the time periods that were used to calculate reference values and lower/higher boundaries (see figure 3).

Boundary values for status classes – MTL 3.25						
Status class	< Bad	Bad < Poor	Poor < Mod.	Mod. < Good	Good < High	High <
Percentile-Approach	3.25	3.84	3.93	4.00	4.10	5
	Lower theoretical	Median value 1997-2000	10thperc. 2009-2014	50thperc. 2009-2014	90th perc. 2009-2014	Higher theoretical
SD-Approach	3.25	3.84	3.92	4.01	4.09	5
	Lower theoretical	Mean value 1997-2000	Mean value – SD 2009-2014	Mean value 2009-2014	Mean value + SD 2009-2014	Higher theoretical

Table 2: Boundaries between the status classes for "bad", "poor", "moderate" (Mod.), "good" and "high" status, calculated based on the Bay of Biscay average yearly Mean Trophic Level 3.25 (MTL_3.25) (table 1). These status classes are parts of the demanded input information in the NEAT tool. For comparison, two approaches were used for calculating the boundaries: by using the median and percentiles (perc. approach), or by using mean and standard deviation (SD-approach). See Fig. 3 for a graphical representation of the percentile method.

Cantabrian Sea in southern Bay of Biscay case study (ES)

In this case, MTL values have been calculated using two data sources: survey data (DEMERSALES survey) from the IEO (Instituto Español of Oceanografía) Database for the Cantabrian sea in Spain (ICES region VIIIc), and ICES DATRAS data for the same region, which are based on the DEMERSALES survey but run along a shorter time period (Fig. 4). The purpose of using both data sets to conduct the assessments was to assess how the choice of specific data sources (comprising shorter or longer time series) may affect the outcome of the analyses, and thus the definition of reference values and boundaries.

The demersal system of the Cantabrian Sea was experiencing a strong fishing pressure during the 1980s-1990s, as reflected by the lower MTLs observed in the beginning of the time series. In response to increased management measures and regulations put into force during the last decade (ICES 2013), the fishing pressure decreased. Since then, the system shows signs of recovery (Fig. 4).

Two approaches were tested for this case study. Firstly, during initial discussions, the EcApRHA Food Webs expert group, agreed to calculate the reference value as the median value between 1992 and 1998, representing the minimum acceptable MTL for this system (reference value of 3.89). This value was used in combination with those of 3.5 and 5 as the minimum and maximum boundary levels accordingly, whilst letting the in-between classes be defined by the NEAT tool. Secondly, after further discussions it was agreed to use a similar precautionary approach as for the northern Bay of Biscay, basing the reference value on the median MTL value during the later time period 2001-2015 (reference value of 3.87). The boundary between "poor" and "moderate" was based on the 10th percentile and that between "good" and "high" on the 90th percentile. Also, as in the first case study, the minimum and maximum range was set to 3.25 and 5, respectively, and the in-between classes to be calculated by the NEAT tool (table 4).

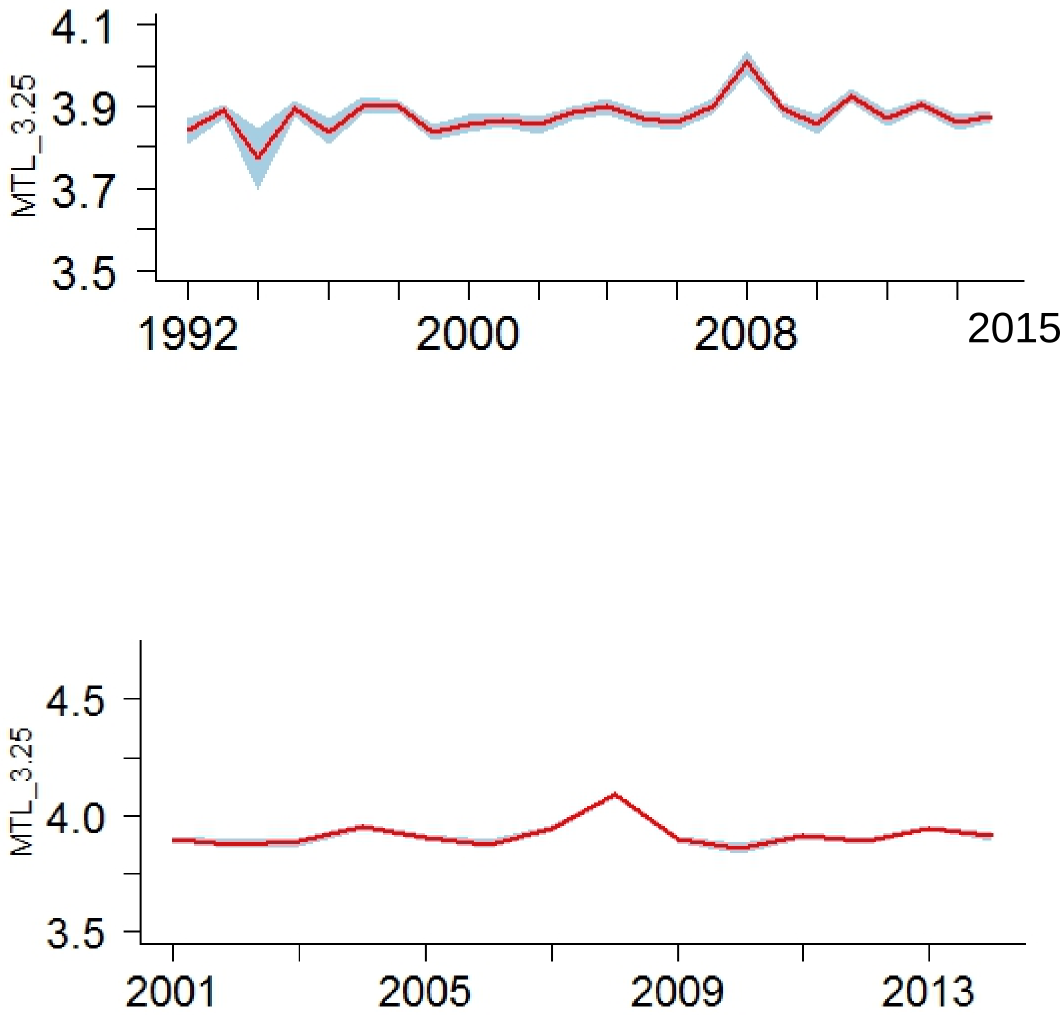


Figure 4: Mean Trophic Level (MTL) calculated from survey data from the Cantabrian sea in Spain, using (a) the IEO Demersales dataset, during the period 1992 to 2015 and (b) the ICES DATRAS Database, for the period 2001 - 2014.

Boundary values for status classes – MTL 3.25						
Status class	< Bad	Bad < Poor	Poor < Mod.	Mod. < Good	Good < High	High <
DATRAS Percentile-Approach	3.25	3.56	3.88	3.90	3.95	5
	Lower theoretical	Intermediate	10th perc. 2001-2014	Median value 2001-2014	90th perc. 2001-2014	Higher theoretical
IEO Demersales Percentile-Approach	3.25	3.89/3.55	3.86	3.87	3.92	5
	Lower theoretical	Median value 1992-1998	10thperc. 2010-2015	Median value 2010-2015	90th perc. 2010-2015	Higher theoretical

Table 4: Boundaries between the status classes for “bad”, “poor”, “moderate” (Mod.), “good” and “high” status, calculated based on the Cantabrian Sea average yearly Mean Trophic Level 3.25 (MTL_3.25). These status classes are part of the demanded input information in the NEAT tool. DATRAS data were used in this table. The two values corresponding to the bad<poor level in IEO Demersales data correspond to the ones calculated directly on the data during years of high fishing pressure, and calculated by the NEAT tool as the mean between the lower theoretical and poor<mod status values, respectively.

The results obtained show that even if the DATRAS series was much shorter than the IEO one, if boundaries are to be set based on the latter 6 years (corresponding to an MSFD cycle in which the trend is a positive/increasing one), as a precautionary approach, then the boundaries between status level remain more or less the same regardless of the data source used. The bad-poor boundary level calculated based on the “bad” IEO years (1992 – 1998) was much higher than the one calculated by NEAT (3.89 vs 3.56, table 4), and even higher than the reference value established for the latter period (2010 – 2015). These results pose questions regarding whether boundary values established by the NEAT programme may be too conservative and not realistic (a value of 3.56 signalling a bad/poor environmental status is not likely to be reached even under high fishing pressure), and highlight the need to include expert knowledge when interpreting/setting these boundaries. The IEO data showed how during years in which high fishing pressure was still on-going, MTL values did not show a drastic decrease, except for a steep drop in 1994 and a slighter one in 1996. These low values were mainly caused by an abnormally high biomass (ca. 30% and 11%, respectively) of *Macroramphosus scolopax*, a species with a relatively low trophic level (TL = 3.5±0.4), and hence the MTL values obtained those years (MTL 3.25 of 3.77 and 3.83) are not necessarily representative of the period as a whole. Nevertheless, they do show that even under abnormal conditions the MTL will not reach levels down to 3.55 as calculated by the NEAT tool and that more realistic values should be adopted if coherent conclusions are to be extracted from the use of the tool. As a result of these considerations, maybe the wisest decision would be to use the 3.77 level value obtained during 1994 as the bad-poor boundary and be ready to closely examine any value coming close to that limit in future assessments. In any case, regardless the value that is set as a reference and the boundaries established, the conclusion should be that all those agreements should be re-examined in each MSFD cycle.

The relatively high levels obtained in 2008 (reflected both by IEO and DATRAS data), on the other hand, were caused by the almost complete absence of blue whiting (normally representing the biggest share of the overall biomass), and the higher relative importance of hake and other high trophic level species during that year.

These results as a whole highlight the need to have a very good knowledge of the species and trends in their abundance/biomass driving the indicator, and whether they are caused by fisheries pressure or not.

Abnormal years such as the ones found in our data, which are not directly related to the pressure under analysis should not be used as reference levels, a mean value being more suited in this case.

Primary production case studies

Sylt-Rømø bight case study

In this case study the PP measurements come from the Sylt-Rømø bight along the north-west coast of Germany, sampled between 1997-2014 (Fig. 5). This specific region is known for having experienced low human impact during the sampled period, and can therefore be considered as a "more" pristine region. Consequently, the primary production is fairly stable during the period sampled.

The expert groups have concluded that the dataset is well suited for estimating the lower and higher reference value in this region. So far, however, they have not concluded exactly how this should be calculated. For the purpose of this study, because the pressures on the bight are low, we defined the boundary values based on the 0, 10th, 25th, 50th, 75th, 90th and 100 percentile of the complete dataset to define boundaries between "high", "good" and "moderate" values (Fig. 5). For the "poor" and "bad" status, literature and theoretical values were chosen.

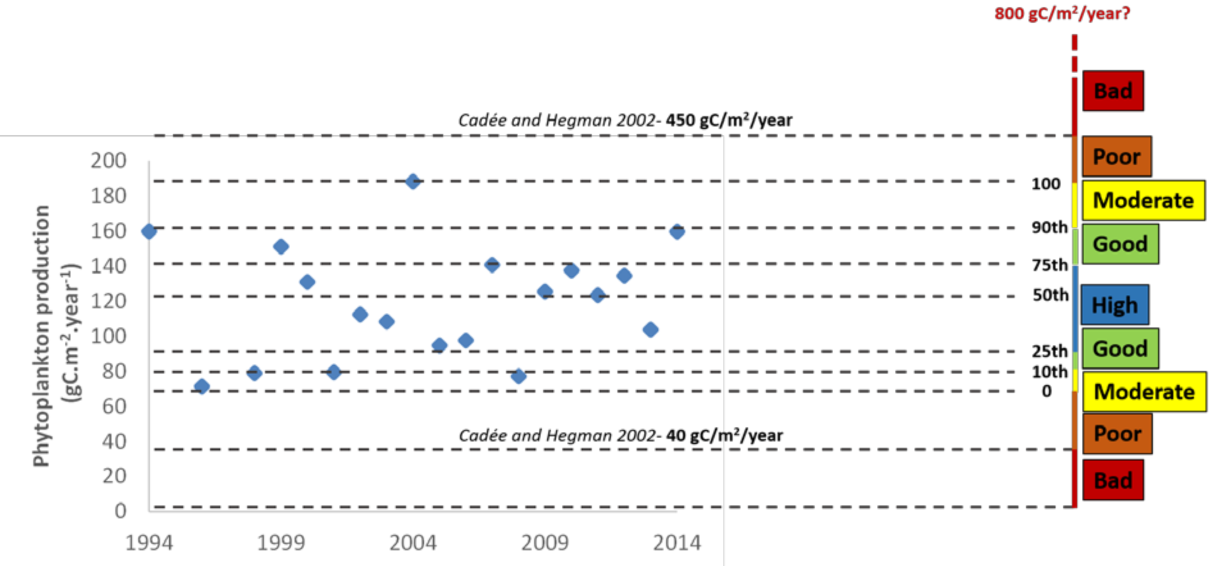


Figure 5: Annual gross primary production (in gC m^{-2}.$year^{-1}$) during the period 1994-2014 from the Sylt-Rømø bight along the west coast of Germany (North Sea). The measurements are based on the oxygen method and are a contribution from Ragnhild Asmus (Alfred Wegener Institute, Germany).The figure illustrates the boundary values for the status classification in NEAT, calculated on the complete time series. For this two-reference value indicator, the 10th, 25th, 75th, and 90th percentile was used, as well as the lowest and highest recorded value. The upper boundary, 450 gC/m2/y, is taken from Cadee and Hegeman (2002), which was the highest value observed as a result of eutrophication according to these authors, whereas the lowest value observed in these time series was 40 gC/m2/year, and we use this value as the lower boundary.

Boundary values for status classes – PP										
Status class	< Bad	Bad < Poor	Poor < Mod.	Mod. < Good	Good < High	High < Good	Good < Mod.	Mod. < Poor	Poor < Bad	Bad <
Percentile	0.00	40.00	71.22	78.40	96.01	138.99	159.49	188.02	450.00	800
approac	Lower	Inter	Lowes	10th	25th	75th	90th	Highe	Inter	High

h	theo.	m.th eo	t value meas.	perc.	perc.	perc.	perc.	st value meas.	mthe o	er theo.

Table 5: Boundaries between the status classes for "bad", "poor", "moderate" (Mod.), "good" and "high" status, calculated based on the complete time series of the Sylt Rømø bight (DE) annual growth primary production. These classes are the proposed boundary classes for indicators with two reference values (i.e., a lower and a higher reference value). Theo. stands for theoretical, interm. for intermediate, and perc. for percentile. See Fig. 4 for a graphical representation.

Bay of Biscay case study

The PP data comes from modelling estimates from the Bay of Biscay from 1900 to 2005 (Fig. 6). The simulated data is obtained from the Coupled Model Intercomparison Project 5, where we are using the average of 8 different scenarios as described in the study (Bopp et al. 2013). The estimated values are given in Net Primary Production, which we transformed to Gross Primary Production by multiplying with a factor of 1.55 (Vézina and Platt 1988; Breed et al. 2004).

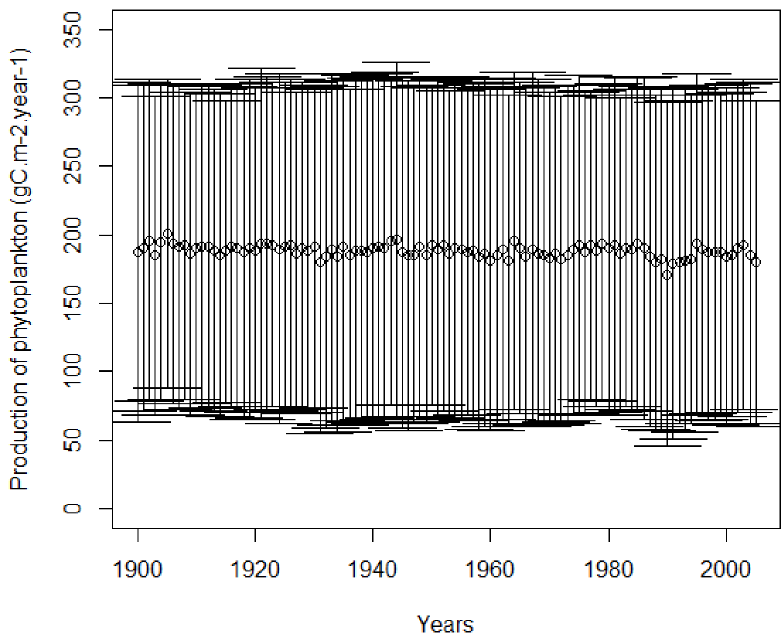


Figure 6: Annual growth primary production (in gC m^{-2}.year^{-1}) during the period 1900-2005 from the Bay of Biscay. The measurements are model estimates from the Coupled Model Intercomparison Project 5, and an average from 8 different environmental scenarios (Bopp et al. 2013). The large spread in the data is given because of the large seasonal differences in primary production

For the purpose of this study, since the data show no significant trend during the period, we used the complete dataset to calculate the boundary levels. We defined the boundary values using the percentiles, the minimum and maximum level recorded, and a lower and upper theoretical value (table 6), as in the Sylt-Rømø bight case study (section 2.2.1).

However, we believe that more data is needed to complete this case study, either estimated in the field or from regional models, and not from regional outputs from global models as those used in the present case.

Boundary values for status classes – PP										
Status class	< Bad	Bad < Poor	Poor < Mod.	Mod. < Good	Good < High	High < Good	Good < Mod.	Mod. < Poor	Poor < Bad	Bad <
Percentile Approach	0	6.5	13	49	82	311	363	397	599	800
	Lower theo.	Interm.	Min. value	10th perc.	25th perc.	75th perc.	90th perc.	Max. value	Interm.	Higher theo.

Table 6: Boundaries between the status classes for "bad", "poor", "moderate" (Mod.), "good" and "high" status, calculated based on the complete time series of the Bay of Biscay yearly gross Primary Production (PP) from 1900 to 2005. The PP values are calculated from modelling simulations and originates from Bopp et al. (2013). Theo. stands for theoretical, interm. for intermediate, and perc. for percentile.

Defining phytoplankton primary production boundaries (including reference value) based on the literature

As for any indicator, defining boundaries is a first step before being able to include the PP indicator in the NEAT tool. However, without boundary levels the indicator itself can be very useful to detect trends. Boundaries can be set from a case study where the data series is sufficiently long to be able to describe PP in the appropriate spatial assessment unit, with the necessity that this time-series covers at least a smaller period of low human pressure (e.g., Sylt Rømø-Bight case study in previous section). However, this is not possible in all case studies, which motivated the use of literature information in order to define the boundary values (Fig. 7).

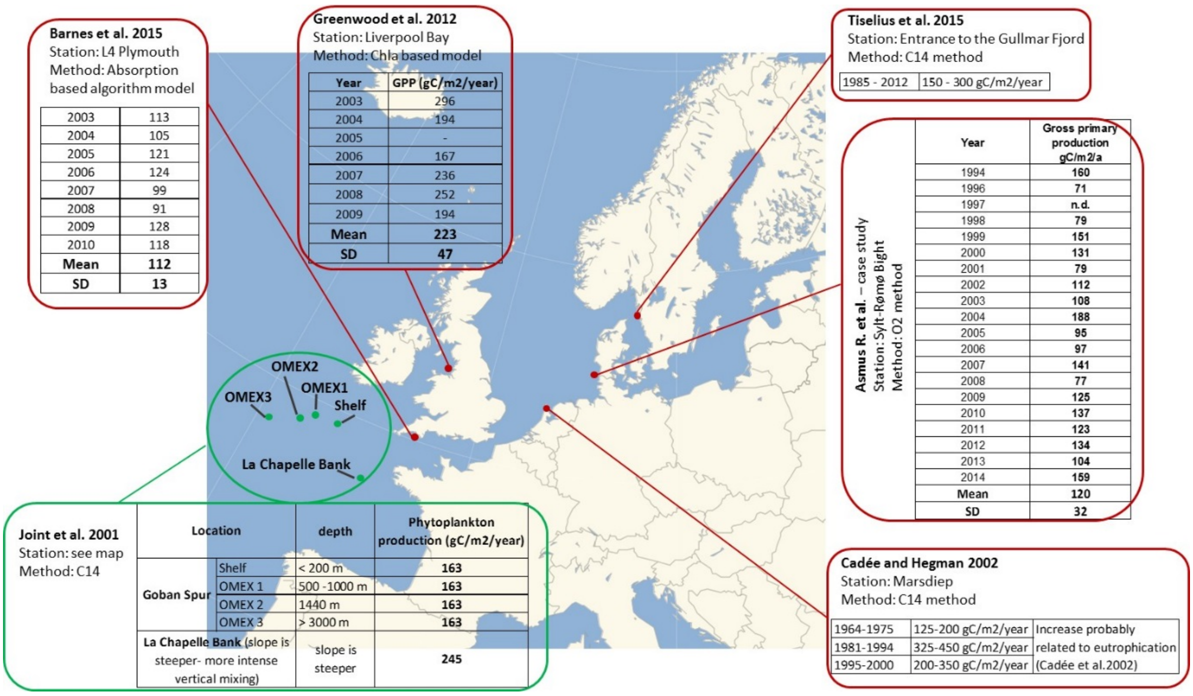


Figure 7: Synthesis of phytoplankton primary production values (in gCm-2year-1) found in the European North East Atlantic waters.

Several authors (Barnes et al., 2015, Cadée and Hegeman, 2002, Joint et al., 2001, Tiselius et al., 2015, Asmus et al. Case study *pers. com.*) observed phytoplankton primary production values mainly in coastal waters (except the OMEX points) varying between 100 and 300 $gCm^{-2}year^{-1}$ (Fig. 7). Cadée and Hegeman (2002) described the PP in Marsdiep (Netherlands) on a very long time-series (1964-2000). These authors observed 3 periods with different levels of PP: (*i*) from 1964-1975, the PP was between 125-200 gC m^{-2} $year^{-1}$, (*ii*) from 1981-1994, the PP was between 325-450 gC m^{-2} $year^{-1}$, and (*iii*) from 1995-2000, the PP was between 200-350 gC m^{-2} $year^{-1}$. In their work, they suggested a relationship between the high increase of PP in the late 1970's and human-induced eutrophication, and the decrease in the latest years due to de-eutrophication. Thus based on that, the high limits observed by these authors (300-450 gC m^{-2} $year^{-1}$) represented the high eutrophic zone in the boundaries definition.

In order to define coastal boundaries characterising the status classes, the values of all studies were gathered (Fig. 7), excluding only the high eutrophic period of Cadée and Hegeman (2002) (*i.e* period between 1981 and 1994) (table 7). By considering all values to define boundaries involved also the mixing of different PP methods (*i.e.,* C14, oxygen, absorption based models and chla based models) should be considered. It is very likely that our approach resulted in very wide boundaries, which may need to be readjusted and be made more regional specific in future monitoring of PP.

Table7: Primary production values (in gCm-2year-1) used to define boundaries. Boundaries were based on the 10th, 25th, 75th, and 90thpercentiles of the data distribution. The highest and lowest values observed are also included (Fig. 8). It should be emphasised that these boundaries are most likely appropriate for coastal systems and may need to be redefined for more off shore and oceanic areas.

Authors	Years	PP (gC/m2/year)
Barnes et al. 2015	2003	113
	2004	105
	2005	121
	2006	124
	2007	99
	2008	91
	2009	128
	2010	118
Joint et al. 2007	1993-1995	163
	1993-1995	163
	1993-1995	163
	1993-1995	163
	1993-1995	245
Asmus et al. Case study (AWI)	1994	160
	1996	71
	1997	-
	1998	79
	1999	151
	2000	131
	2001	79
	2002	112
	2003	108
	2004	188
	2005	95
	2006	97
	2007	141
	2008	77
	2009	125
	2010	137
	2011	123
	2012	134
	2013	104
	2014	159
Greenwood et al. 2012	2003	296
	2004	194
	2005	-
	2006	167
	2007	236
	2008	252
	2009	194
Cadée and Hegman. 2002	1964-1975	125
		200
Tiselius et al. 2015	1985-2012	150
		300

Percentile	
10%	91
25%	108
50%	136
75%	165
90%	239

minimum value

maximum value

The various boundaries (Fig. 8) are defined from the following percentiles (Fig. 8):

The centre of the distribution is the 50thpercentile (*i.e.* 136 gC m^{-2} $year^{-1}$) (note, this is not a direct boundary in NEAT)

The "high" class is limited by the 25thand 75thpercentile boundaries

The "good" class is limited by the 10thand 90thpercentile boundaries

The "moderate" class is limited by the lowest and the highest values observed in the data distributions

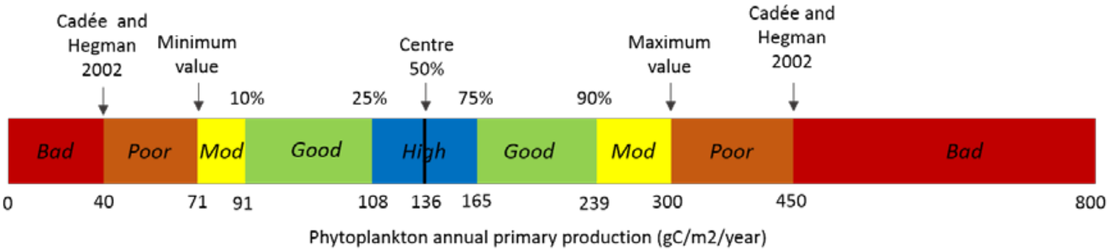


Figure 8: Boundaries of phytoplankton annual primary production in European waters based on literature (in gC m-2 year-1). We want to highlight that the boundaries proposed here are not ideal. They were proposed in this report for the purpose of using the PP indicator in the NEAT tool, and should not be used without careful consideration. Local reference values should be defined considering local drivers of the phytoplankton primary production.

The high "poor" class was based on Cadée and Hegeman (2002) corresponding to high eutrophic boundaries observed by these authors (*i.e.* 300-450 gC m^{-2} $year^{-1}$)

The low "poor" class was also defined from data from Cadée and Hegeman (2002). Indeed, these authors omitted from their data series the earliest data from Postma (1954) because they considered the values (*i.e.* 20-40 gC m^{-2} $year^{-1}$) as obviously too low. From this information, the value of 40 gC m^{-2} $year^{-1}$ was used as the low boundaries for the poor class.

The highest limit of the high "bad" class (*i.e.* 800 gC m^{-2} $year^{-1}$) was taken from the North Sea modelling case study (see section 2.2.2.3). The lowest value for "bad" class was arbitrarily set at 0 gC m^{-2} $year^{-1}$.

We want to highlight that the boundaries proposed here are not ideal. They were proposed in this report for the purpose of using the PP indicator in the NEAT tool, and should not be used without careful consideration. It is important for future assessments to define local boundaries for each case study in order to refine the class boundaries according to local specifications and human-induced pressures.

In the following sections we review several case studies where the definition of the boundaries is difficult to establish due to limited data and/or information. Instead of trying to define their specific boundaries, we are interpreting the data given the "global" boundary levels given above.

<u>Oosterschelde estuary case study for the PP</u>

Phytoplankton PP measurements for this case study come from the Oosterschelde estuary in the Netherlands and derive from biweekly ^{14}C incubations and interpolations by using irradiance, data (Fig. 9). Since 1987, the water exchange between the estuary and the North Sea has been reduced due to a large-scale coastal engineering project, with reduced inflowing water as a consequence (I.e. the residence time changed from 5 days to 10 days). The decreased PP observed (Fig. 9) is hypothesized to be caused mainly by overgrazing by bivalve filter feeders: mainly mussels, which are cultured in large numbers in the estuary and large quantities of "wild" oysters and perhaps *Ensisdirectus* (Smaal *et al*. 2013).

This case study illustrates an example of a decreasing PP level with increasing pressure, indicating the need for a lower reference value. Because of the large external changes to the system, and the persistent grazing pressure from bivalve aquaculture and wild oysters, the expert group has so far not found a solution in how to define the reference value.

The global PP boundary levels defined from the literature (Fig. 8), indicate that the system was in a "moderate" to "bad" environmental status during the initial period from approximately 1991 to 1996, possible a reflection of the eutrophication of the coastal sea also seen in the data from Cadée and Hegeman (2002), which decreased with time. The system then entered a period of "good" to "high" environmental status between 1997 to 2006 or 2012, finally, due to the grazing pressure by filter feeders, the system seems to have entered a critical "moderate" to "bad" level again in 2013. In this case the "bad" level is the result of a too high (we hypothesize) grazing rate which likely will also negatively affect the economic value of the mussels.

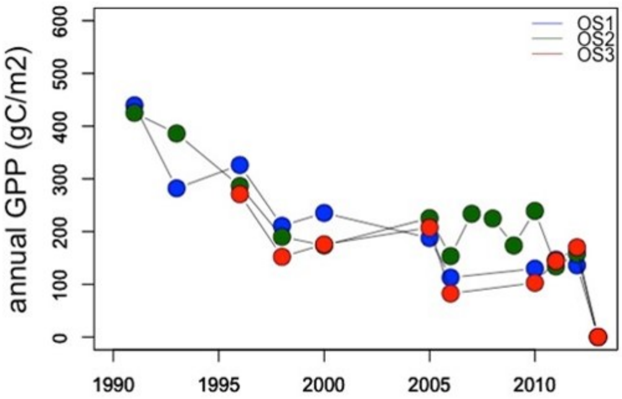


Figure 9: Annual gross primary production (in gC m-2 year-1) during the period 1991-2013 from the Oosterschelde estuary along the coast of Netherlands (North Sea). The measurements are based on the primary production model calibrated with biweekly C14 incubation measurements. The data originates from Jacco Kromkamp (Royal Netherlands Institute for Sea Research, Netherlands).

English Channel case study

For the English Chanel, PP data is only available for a few discrete years (Fig. 10). This data is clearly too short to define any reference value or boundary levels.

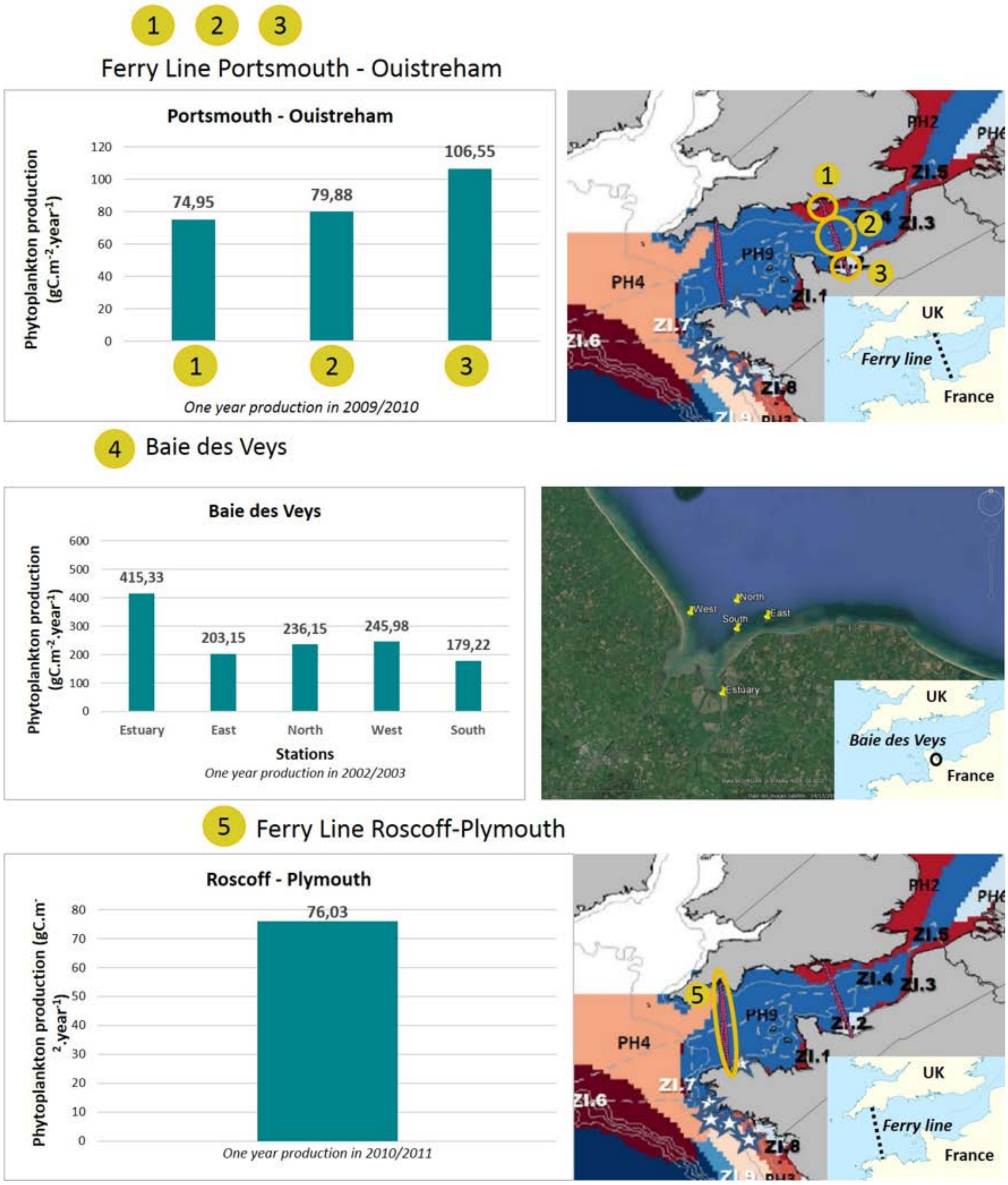


Figure 10: Primary production (in gCm-2.year-1) in different water masses in the English Channel. One year of data were surveyed in three different periods of time: 2002/2003 in Baie des Veys (point 4) using 14C method. 2009/2010 on a ferry line between Ouistreham and Portsmouth (points 1, 2 and 3) using the electron transport rate method with a Pulse Amplitude Modulated flourometer (PAM). 2010/2011 on a ferry line between Roscoff and Plymouth (point 5), again using PAM. Hydrodynamic water masses were taken from Gailhard-Rocher et al., 2012. The data is a contribution by Pascal Claquin from the University of Caen (France).

When the data were compared to global PP boundary levels (Fig. 8), the different measures or case studies were found to be in a "good" to "moderate" status. Data were distributed between 75gC m^{-2} $year^{-1}$ and 246gC m^{-2} $year^{-1}$. Only one exception was found for the estuarine zone of Baie des Veys with a high value of

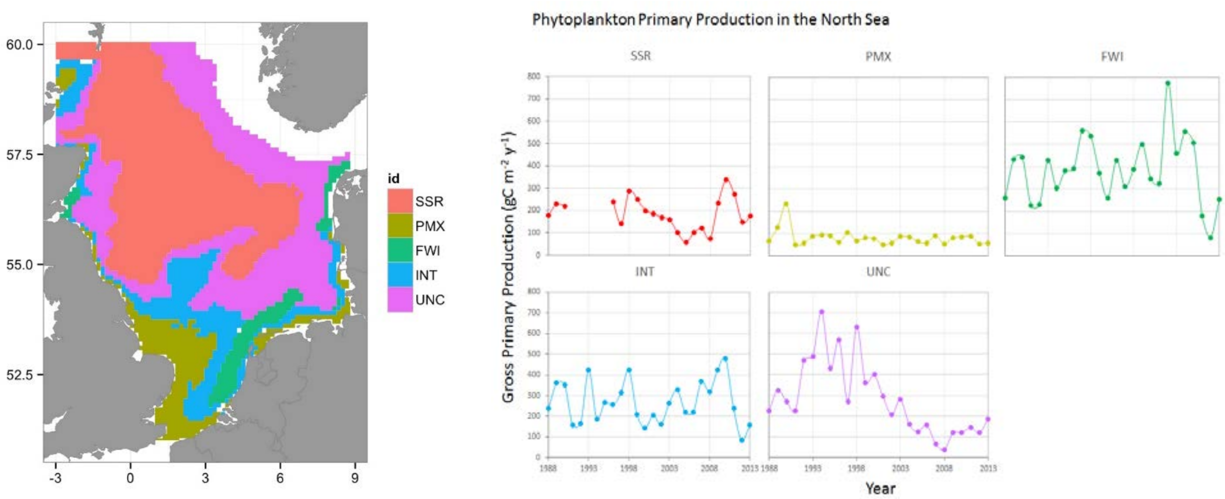
415 gC m^{-2} year^{-1}. The estuarine zone seems in a "bad" status with high PP production and this could be related to high eutrophication in this area

North Sea case study

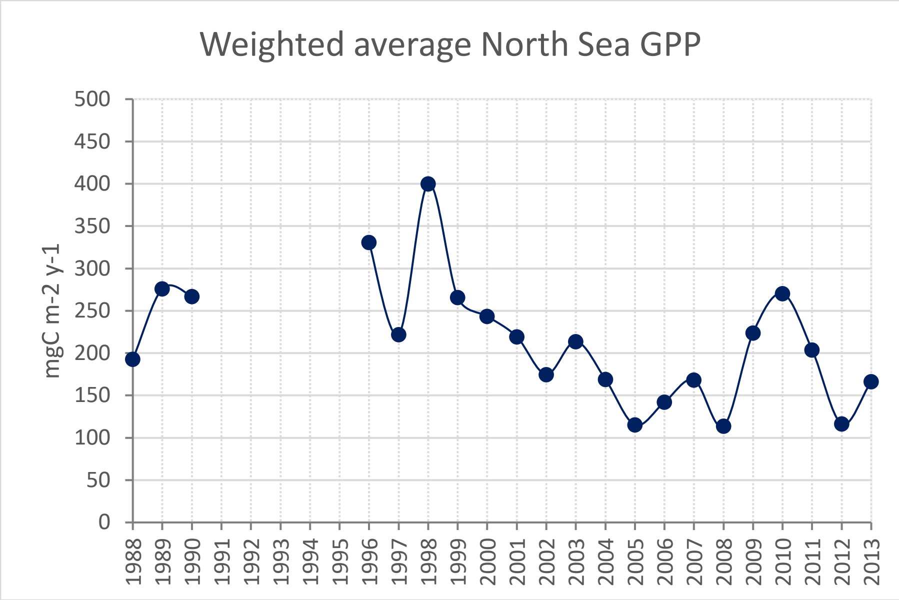
The PP indicator for the wider North Sea was estimated based on eco-hydrodynamic regions (van Leeuwen *et al*. 2015; Fig. 11a). PP was calculated adopting a semi-empirical algorithm (Cole and Cloern, 1987), based on measurements of chlorophyll concentration, the integrated daily irradiance and underwater light climate (depth of the photic zone), which were grouped by eco-hydrodynamic regions. In this case study, we provide an overview of the PP time series at the different eco-hydrodynamic regions (Fig. 11b), and of the weighted average PP of the wider North Sea (Fig. 11c).

The regions are characterized by different areas, hydrodynamic conditions and nutrient regimes; therefore, PP was highly variable between years and between regions. The offshore region (SSR) and the coastal region off East Anglia (PMX) showed the lowest PP (approximately 50-200 gC m^{-2} year^{-1}; Fig. 11b). Compared to the global PP boundaries (Fig. 8), the PP in these regions have been fluctuating between "low moderate" and "high moderate", with periods of "good" and "high". It is important to highlight that production in these regions is low for different reasons. SSR includes clear offshore waters, where production is generally limited by nutrient concentration. Contrarily, the PMX region is characterised by higher nutrient concentration but the high turbidity of the water column (as result of high concentration of suspended sediments) limits phytoplankton production.

The other regions had higher PP values (approximately 200-500 gC m^{-2} year^{-1}), with occasional peaks above 500 gC m^{-2} year^{-1}. Coastal areas are typically subjected to larger nutrient loadings compared to offshore areas, and can show higher values of PP. The weighted average of the wider North Sea (Fig. 11c), which combined production of the different regions weighted by their area, varied between 100-300 gC m^{-2} year^{-1} circa, equivalent to a "high/good" to "moderate" status (compare with scale in Fig. 8).



a) b)



c)

Figure 11: Eco-hydrodynamic regions of the North Sea (van Leeuwen et al. 2015) (a), and gross primary production (in gC m-2 y-1) from 1988 to 2013 in the different eco-hydrodynamic regions (b), and as weighted average of the wider North Sea (c). PP was calculated adopting a semi-empirical algorithm (Cole and Cloern, 1987), based on measurements of chlorophyll and underwater light climate, and is an extract from Capuzzo et al. (manuscript in preparation) .SSR: seasonally stratified; PMX, permanently mixed; FWI: Freshwater Influenced, INT, intermittently stratified, UNC: unclassified. The gap between 1991 and 1996 is because no light data were available for the seasonally stratified region (SSR).

4 NEAT assessment

In this section we illustrate how several food webs indicators can be integrated into the NEAT tool for a complete assessment. For this purpose, we are using the Bay of Biscay case study, using data and indicator reference values and boundaries defined for MTL 3.25 and PP. We are illustrating the application of NEAT by showing the main steps in the software (**Annex 1**).

Applying reference values and boundaries for the Bay of Biscay case study

The information needed for doing an assessment are the reference and boundary values (**table 8 below**) and the assessment values (**table 9 below**) for each indicator.

For the purpose of illustrating a complete NEAT assessment, we had to modify the boundary levels for the indicators with two reference values (i.e. PP) (**table 3 and table 6**).

Copy of table 6:

Boundary values for status classes – PP										
Status class	< Bad	Bad < Poor	Poor < Mod.	Mod. < Good	Good < High	High < Good	Good < Mod.	Mod. < Poor	Poor < Bad	Bad <
Percentile Approach	0 Lower theo.	6.5 Interm.	13 Min. value	49 10th perc.	82 25th perc.	311 75th perc.	363 90th perc.	397 Max. value	599 Interm.	800 Higher theo.

Since the current version of NEAT only allows for 5 classes (i.e. “Bad”, “Poor”, “Moderate”, “Good”, “High”) on a monotonic scale, we had to reduce the proposed 9 classes (see copy of **tables 6**). This was done by using only the upper 5 classes from "high" to "bad", since the indicator values were within the higher range (table 9). Further, in order to allow for the classes to go from low values representing "**bad**" status, to high values representing "high" status, the boundary levels and indicator values were reversed by multiplying them by -1. This type of modification is not desirable, and will not be needed in forthcoming versions of the NEAT tool.

Boundary values for status classes for NEAT assessment – Bay of Biscay						
Status class	< Bad	Bad < Poor	Poor < Mod.	Mod. < Good	Good < High	High <
MTL 3.25	3.25	3.84	3.93	4.00	4.10	5
PP (literature)	-800	-450	-300	-239	-165	-108
PP (case study)	-800	-599	-397	-363	-311	-82

Table 8: Boundary levels for the food webs indicators used in the Bay of Biscay NEAT assessment. For PP, only the upper half of the boundary levels, including only the upper reference value, were used in the NEAT assessment since the current version of NEAT only allows for 5-status classes. The negative value of these boundary levels has been used to allow the right order of status classes ("bad" as the lowest value, "high" as the highest values, etc). In future assessment work, however, suggestively the full range of boundary levels should be used, including both the upper and lower reference values. The detailed tables with information for the different indicators can be found in table 2 for Mean Trophic Level (MTL) 3.25, table 6 for Primary Production (PP) Bay of Biscay case study, and Fig. 8 for Primary Production (PP) literature values.

.

Indicator assessment values ± SE for NEAT – Bay of Biscay				
	MTL 3.25		PP	
	Mean	SE	Mean	SE
6 year interval				
1997-2002	3.86	0.01	-193	17
2000-2005	3.91	0.02	-192	17
3 year interval				
1997-1999	3.84	0.02	-193	24
2000-2002	3.87	0.02	-192	24
2003-2005	3.94	0.02	-192	24

Table 9: Average indicator values with associated Standard Error (SE) for Mean Trophic Level (MTL) 3.25, and Primary Production (PP) for a 6 and 3-year period (These are random periods for assessment purpose). These are the assessment values used in the NEAT assessment. Note that these values were taken from the same dataset as where the boundary levels were defined from. This was done for illustration purpose only

For this assessment we used the average indicator value for 3 or 6 year's intervals (these are random periods for assessment purpose).

Assessment of Bay of Biscay environmental Status

We are here presenting a synthesis of the work-flow in NEAT using one of our case studies. This is not complete in any sense, and we therefore recommend the user to turn to the NEAT work manual and to see the NEAT webinar for detailed information (http://www.devotes-project.eu/neat/).

Three major steps were followed in order to prepare the assessment run using the Bay of Biscay case study (For more details, see Annex 1).

1) The first step was to define the Spatial Assessment Unit (SAU), the habitat and the ecosystem components.

2) The second step was to define indicators and set the boundary levels of each indicator.

After having entered or modified all indicators, it is time to pick out the indicators for each spatial assessment unit in the assessment

3) The third step is to evaluate the result, and include potential weighting-factors.

The result from our assessment is show in Fig. 12. Based on our two food webs indicators, the Bay of Biscay is in Good Environmental Status. The two different versions of the PP boundary levels used (from literature and a case study, respectively), show that defining the boundary levels from the case study is slightly more conservative, and the status value is slightly lower. We can also observe a slight increase in the status value during the last years (Fig. 12).

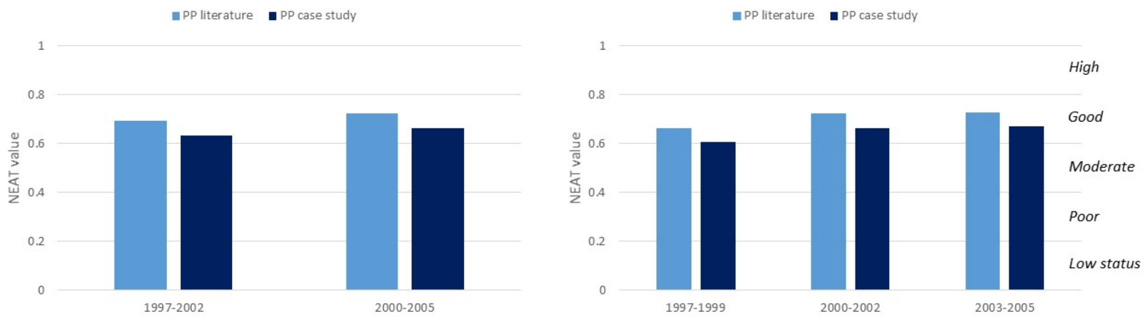


Figure 12: Results of the NEAT assessment for 6 year intervals (left panel) and 3 years intervals (right panel). The assessment was run on two versions of Primary Production (PP) boundary levels. Those defined from a literature review (PP literature), and those defined from the local case study (PP case study). The lines indicate the boundary levels for the status classes.

BENEFITS AND CHALLENGES WITH NEAT TOOL

The main benefits of testing the NEAT tool were:

- Linking the food web indicators (and potentially indicators from other ecosystem components) toward a more holistic assessment of regional and sub-regional Good Environmental Status of marine ecosystems;

- Identifying technical and ecological limitations for including indicators in NEAT. The technical limitations should be addressed in future NEAT versions;

- Testing different scenarios. In this report, we tested the influence of setting different reference values (i.e. for FW2 indicator, setting local reference values vs. literature values) on the final assessment result. The tool allows to test much more scenarios and to apply a sensitivity analysis on assessments. However, it was not feasible to test all the possible scenarios within the EcApRHA project.

The main limits of our testing:

- Only two food web indicators were tested in this report, which limited the possibility of testing a more complex nested approach with different spatial assessment units;

- Reference values that were tested in this report are based on local experts' knowledge of their ecosystems for the two food web indicators used here. No food web values have been agreed between OSPAR Contracting Parties. Setting reference values is needed for NEAT tool (in its current version) to be able to run an assessment and this can be an important limitation for including indicators.

GUIDANCE FOR FUTURE WORK

The NEAT software is preconfigured for use in the four European Regional Seas with their assigned indicators per Texeira et al. (2014). If the tool is sufficiently relevant for Contracting Parties, its applicability for assessing regional and subregional marine environmental status in OSPAR regions can be quickly done.

This report is a first attempt to test the NEAT tool for integration of food webs indicators. However, even if the tests were limited here, we do think that the NEAT tool should be further considered for use within OSPAR. A timeline of actions (Fig. A) is proposed below which would allow more testing and scenarios production with NEAT in upcoming years (additional information can be found in Elliott et al., 2017 and Padegimas et al., 2017).

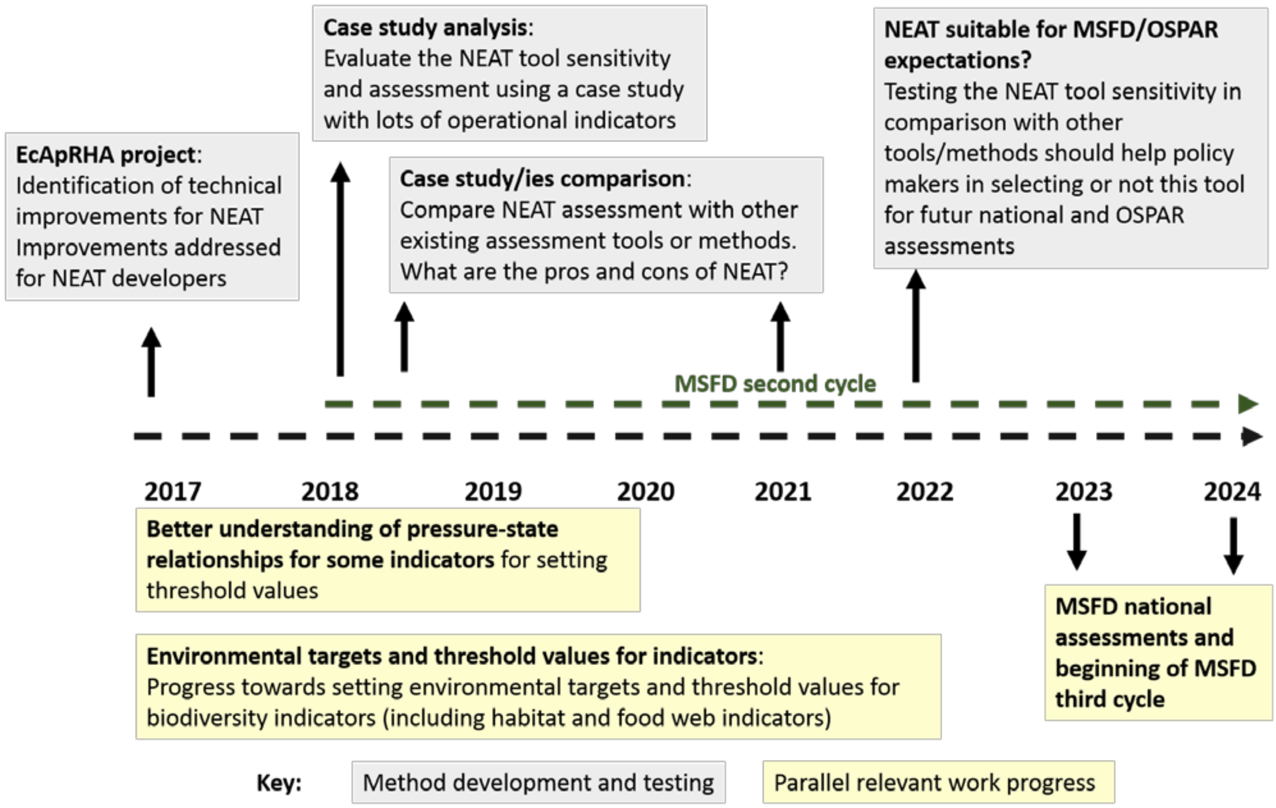


Figure A: Timeline along which food web integration using the NEAT tool could be additionally tested and implemented with the assumption of sufficient resource

5 Gaps and solutions

Knowledge gaps

This section identifies knowledge gaps towards the use of NEAT tool for assessment purposes. The list of knowledge gaps is not exhaustive but rather based on our experience of applying food webs indicators into NEAT. Some broader ecological questions are also addressed (Padegimas et al., 2017).

Define reference values or a methodology to set these boundaries for all indicators that need to be included in NEAT tool.

One of the critical pieces of information that is requested for the use of NEAT tool is the definition of status classes' boundaries. These boundaries are based on the reference values defined for each indicator. This implies a definition of the reference values but also the adoption of these values by member states towards the use of these boundaries in NEAT within a sub-regional assessment.

The current revision of the Commission Decision (2010/477/EU) indicates that all indicators will need to have assessment thresholds by 2024. This implies the need to define reference values in the upcoming six years (i.e. 2016-2022) in order to meet the requirements of the Commission Decision for the third MSFD assessment in 2024.

Reference values are needed for the NEAT tool to evaluate change of the environmental status. However, constructing them can be challenging, particularly when ecosystem component data availability is limited. However, indicators can also be used in absence of reference/threshold values in trend analyses, which can be quite powerful. But for a combination of different indicators, reference values are required if the NEAT tool is to be applied in its current version. In the current document, a case study approach was used to define local reference values in order to test the integration of food webs indicators in the NEAT tool. A more thorough study using a wider variety of regions needs to be conducted so that reference values can be based on better knowledge of the indicator's fluctuations. Using a case study approach needs to be developed considering a spatial approach, looking at the spatial influence of environmental parameters, and a temporal approach, focusing on change over time.

A precautionary approach should be set with an update of the reference values for each MSFD cycle (i.e. each six years) based on experts knowledge. In the current document, the reference values proposed for the integration in NEAT tool are based on the experts' knowledge of their respective ecosystems. The reference values defined reflect the experts' scientific knowledge of each ecosystem status. However, these values that are proposed are not definitive. A precautionary approach should be applied and the reference values should be updated in each new MSFD cycles based on the experts' knowledge evolution of their ecosystem.

Finally, it is worth noting that the reference values that are proposed in this report are calculated for the purposes of demonstrating the NEAT tool and do not represent any agreement of reference values by European Member States that are OSPAR Contracting Parties.

Integrate more indicators with different scales and different ecosystem components for a nested approach and make a sensitivity analysis

In the current report, the integration of two food webs indicators in the NEAT tool, to make an assessment of the Bay of Biscay continental shelf, used a unique spatial assessment unit for both indicators (i.e. Mean Trophic Level indicator and Phytoplankton production indicator). The test showed how two different indicators (i.e. describing the structure and the functioning of the Bay of Biscay continental shelf food web) could be integrated to evaluate one spatial assessment unit. However, the lack of data at different spatial scales and the low number of indicators that were considered did not allow to have a complete nested approach in the current assessment. The nested approach should be applied in future tests by increasing the number of indicators in a case study area. These indicators should be at different spatial scales.

A sensitivity analysis, which is already available in the current version of the NEAT tool, could be also applied in this case study area to produce various assessment scenarios. The scenarios could investigate, *inter alia*, the influence of:

- increasing or decreasing the number of indicators in the analysis;

- using different time scales/indicator and their influence on the overall assessment;

- applying indicators at different spatial scales and see the influence on the overall assessment;

- separating or aggregating the information of the various MSFD descriptors

These scenarios should be developed considering the management implications and where possible, propose optimisation of MSFD elements (e.g. monitoring programme).

Discriminate human/natural changes effects on the indicators before setting reference values

For some indicators, it is difficult to set reference values and environmental targets that could have management implications without previously disentangling the various pressure effects and further investigate the pressure-state relationships. When changes are detected for such indicators, it is often difficult to discriminate between direct human induced changes in pressure (i.e. anthropogenic pressure) from the natural or indirect human pressure change (e.g. prevailing conditions or climate change). To cope with these challenges, some additional research work should be conducted in the future years to:

- Investigate changes in state indicators from those driven by prevailing conditions (Identify the various pressures influencing each indicator and analyse the pressure-state relationships);

- Indicator responses to key anthropogenic pressures and climate drivers should be quantified to determine magnitude and direction of indicator change;

However, expert evaluation of the results will always be necessary to check the final results.

Methodological gaps

This section summarises the key limitations encountered in the integration of the food webs indicators in the NEAT tool. It aims to address the identified limitations to the designers of the NEAT tool so that they could consider them in future versions of the tool. Thus, the solutions should be proposed by the designers of the NEAT tool. This section will also propose some broader questions not tested in this report.

Develop more flexibility in the current five status classes (i.e. “Bad”, “Poor”, “Moderate”, “Good” and “High”) of the NEAT tool

Testing the integration of the food webs indicators in the NEAT tool has underlined the limitation of applying some indicators, which needs to have higher and lower limits to define the “Good” status of the assessed area, in the current version of the tool. The phytoplankton production can be considered in a “Poor” status if its value exceeds a certain upper limit (e.g. 300 $gC.m^{-2}.year^{-1}$) but also if its value drop under a lower limit (e.g. 70 $gC.m^{-2}.year^{-1}$). Our analysis revealed the necessity of having 9 status classes for such indicator as shown in the Fig. 2.

Develop a confidence around indicators with their combined SAU/Habitat/Ecosystem component

In the current version of NEAT, there is no weighing applied at the indicator level to avoid that an indicator could dominate the assessment result. However, there is a need for a level of confidence to be given for each indicator with its related Spatial Assessment Unit (SAU) and habitat. For example, for some regions we have access to time series and can, based on these, estimate the boundary values for the different status classes. For other regions, however, we may have only a few data points, and we may need to base our boundary values on literature information. Consequently, we will put less confidence in these indicators in these regions, and would like this to be accounted for in the analyses.
In the current version of NEAT it is only possible to include uncertainty in form of a standard error of the assessment value of each indicator. There is today no way to include uncertainty (or degree of confidence) for an indicator for a specific region. This question needs to be addressed by the NEAT tool developers.

Make available the uncertainty of the assessment in another format than the current visual bars

The current version of NEAT gives the final assessment as a “NEAT value” between 0 and 1 and the associated “Uncertainty” is presented in the form of bars distributed in the various status classes (see figure below).

NEAT value	Status class	Uncertainty
0.589	moderate	
0.489	moderate	
0.493	moderate	
0.216	poor	
0.632	good	
0.689	good	
0.161	bad	
0.080	bad	
0.079	bad	
0.773	good	
0.886	high	

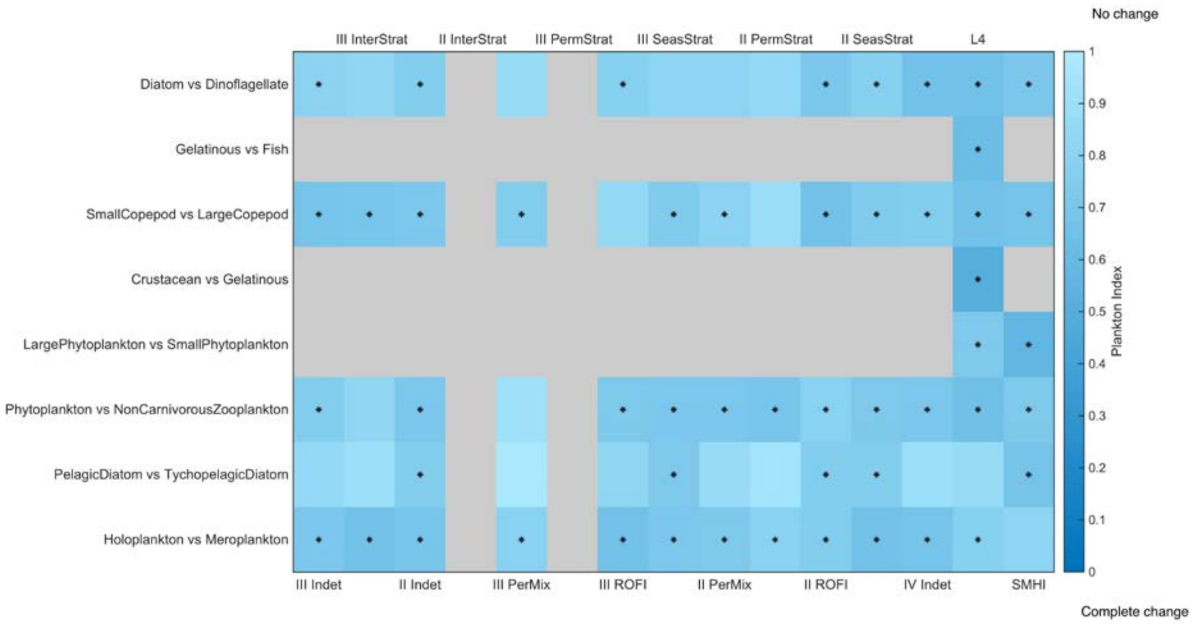
However, when the assessment is extracted (in a csv format for example), only the “Neat value” is extracted and the “Uncertainty” is not available. The forthcoming version should consider the transformation of uncertainty bars into numerical values giving uncertainty to the NEAT value.

Map representation of the assessment

The current version of NEAT gives the final assessment in a table format where we have the final assessment NEAT value (i.e. between 0 and 1 with 0.6 representing the reference value) for each SAU and the nested sub-SAUs and for the different habitats that are embedded. Can these information be represented in a Map format in future version of NEAT?

Can an adaptive scale be developed for use within NEAT?

In order for the NEAT tool to be used with state-space plankton indicator, the current scale of ‘good to bad’ that is used within NEAT would need to be changed to a significance scale (proportion of points outside of the state-space plot). This scale would need to go from 0 – 1, but use a colour scheme that does not represent good/bad, but is neutral. For Example, see Fig. 13 demonstrating the scale used:



Figure

Figure 13: Change in Plankton Index for the period 2009 - 2014 from starting conditions (2004 – 2008) for each lifeform pair. Darker blue indicates a more pronounced change. Grey shading represents where there were not enough/well-represented data to determine a Plankton Index. Starred cells indicate significant change (p<0.01) from starting conditions.

Changes in the Plankton Index do not necessarily indicate a deterioration of environmental conditions; they do, however, indicate change from starting conditions, and give thus an incentive to investigate the reasons behind the change in order to know if the observed change will affect the ecological status assessment in a desired or un-desired way.

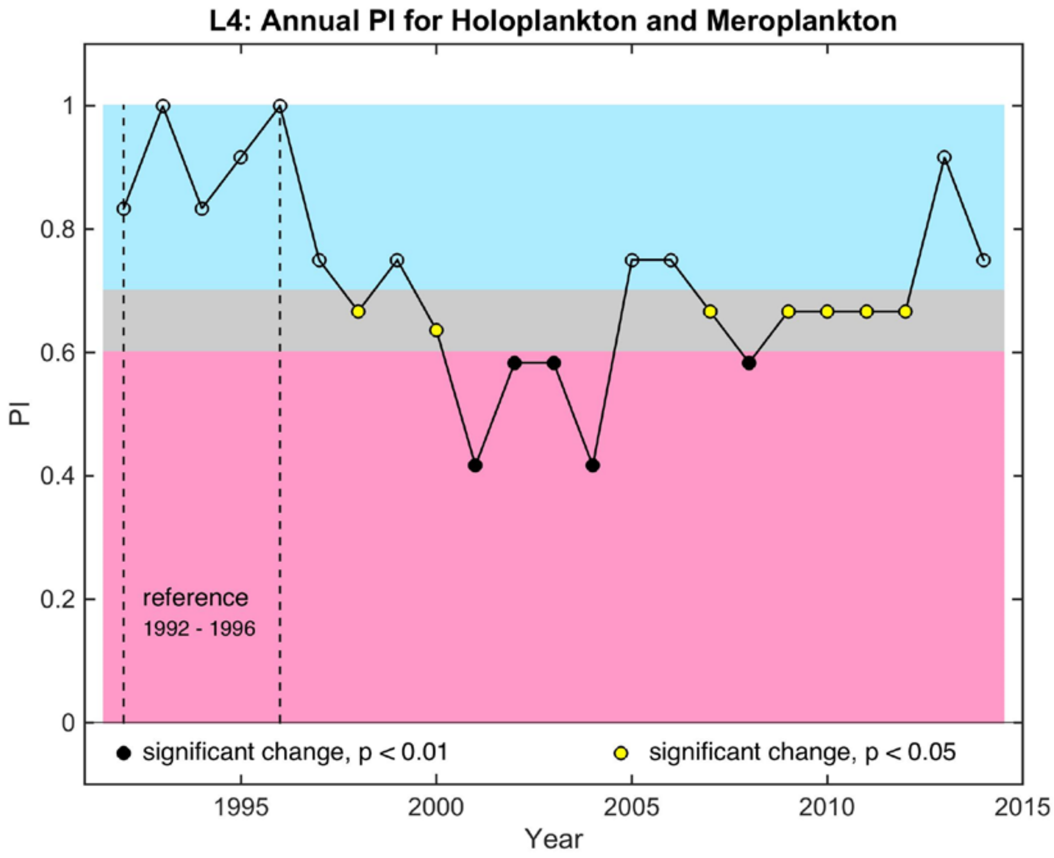


Figure 14: Change in the annual Plankton Index for the lifeform pair Holoplankton and Meroplankton recorded at station L4 from 2004 - 2014, using the starting conditions 2004 – 2008. Light blue area is the threshold for no significant change from the starting conditions, grey is a significant change for $p < 0.05$, and pink is a significant change for $p < 0.01$.

Changes in the annual Plankton Index do not necessarily indicate a deterioration of environmental conditions; they do, however, indicate change from starting conditions.

6 Questions addressed to the NEAT developers

During the work within EcApRHA, several difficulties were addressed within the expert groups in how food webs indicators can be integrated into the NEAT tool. In order to answer some of these questions, we discussed these issues with one of the developers of the NEAT tool, Torsten Berg (MariLim aquatic research GmbH, Germany). Below is a summary of the questions addressed and the potential solutions proposed.

"How can we accommodate our indicators to the NEAT tool, and how can the NEAT tool accommodate to our indicators":

Direct methodological questions**:**

How can other components be included in the NEAT interface, such as habitats or ecological groups? For example, when using a food webs indicator such as Mean Trophic Level, the indicator combines information from both pelagic and benthic habitats, and for several trophic groups (benthic and pelagic organisms, and in the future, also marine mammals and birds).

Also, can "ecosystem" indicators (as described above) be combined with other indicators which belong to specific habitats and/or trophic groups?

Reply: Practically it is possible and easy to include a new component into the NEAT tool, and any name can be assigned to that new component. For a food webs indicator for example, a new habitat such as "ecosystem" can be included, or a new ecological group such as "food webs". (Practically: on the opening window of NEAT, click on the component groups of interest (habitat, ecosystem component etc.). This will open a new window. By pressing on the "+", another new window will open where a new component can be defined)

The SAU and habitats can be entered into NEAT with a nested design (different levels). It may therefore be important to consider at what hierarchy (level) a food webs indicator is in relation to others. For example, if the habitat defined for a food webs indicator (we can suggestively call it the "ecosystem" habitat) should be on a higher hierarchy/level compared to for example a pelagic or benthic indicator. It is important to realise that for an indicator which is linked to a habitat that is assigned a higher level, the assessment NEAT value at this level will be imposed on all other lower levels. This means that this indicator will have strong influence on all other indicators. If this is not desirable, but the hierarchical order between habitats is motivated, this can be circumvented by giving the food web habitat a priority factor <1 (to find an appropriate value, this may need to be tested). In conclusion, priority can be given at the SAU or habitats level but not at the "ecosystem" habitat level.

Note: A priority factor should not be applied if there is no strong reasoning for this. The weighting system in NEAT is separate from the priority factor. The weighting system is defined through the size of the SAU-area. The priority factor can be applied to the SAU or habitat, and should only be used when there are good reasons e.g., when a geographically small area has comparable high importance in the assessment, such as marine protected areas.

How can we deal with an indicator which has two reference values, both a minimum and a maximum level? The Primary Production indicator for example, both a too low and too high level indicate a bad status, meaning low and high reference values are needed.

Reply: There are two alternatives to this issue:

1) The PP, which has a unimodal distribution (in terms of status classes), can be transformed mathematically into a monotonic distribution. Torsten had no direct suggestions, but this will be looked into? This is a possible solution, although not ultimate.

2) The forthcoming version of the NEAT tool will, however, have more flexibility in how to define the status classes, including the possibility to include a lower and higher reference boundaries. This version (version 1.3) was not available at the time this report was being finalised, so its suitability could not be checked.

How to deal with different levels of confidence between the different indicators. Is it possible to include uncertainty around the indicators, which was possible in the previous version of BEAT? For example, for some regions we have access to time series and can, based on these, estimate the boundary values for the different status classes. For other regions, however, we may have only a few data points, and we may need to base our boundary values on literature information. Consequently, we will put less confidence in these indicators in these regions, and would like this to be accounted for in the analyses.
This is an important point from an EcApRHA perspective, since it would open up the possibility to include more indicators (and from more regions) even if the level of confidence differed.

Reply: In the current version of NEAT it is only possible to include uncertainty in form of a standard error of the assessment value of each indicator. There is today no way to include uncertainty (or degree of confidence) for an indicator for a specific region. This would suggestively best be done by including an uncertainty around the boundary levels of the status classes. Technically this is possible to do, and would then contribute to the uncertainty of the result estimated through the Monte Carlo simulation. Torsten Berg will discuss this question with other developers.

How important are the "end values" of the classes (lower value of the bad status, versus higher value of the high status)? For the MTL for example, we consider to use theoretical values that are practically impossible to reach. Will these values influence our results (and the normalized scale)?

Reply: This has not been specifically tested. However, Torsten Berg assumes that this will not have any important effects on the results. E.g., if the boundary value is set to be unreachable, the indicator will always end up close to the next boarder value which will assure that the weighted average operates on a realistic scale.

Note: If an assessment value is outside the outer boundary levels of the status classes, the program will truncate the value to the lower/higher boundary level. However, there will be no warning or information from the software that this has been done.

And, is it important that the boundary values for the status classes are equally spaced (as you suggested in the webinar)?

Reply: No, this is not important. It is the ecological relevance of the boundary values that are important.

Suggestively, it is typically good to start defining the reference value for "Good Environmental Status". From there, other possible meaningful divisions should be considered. If the user is not willing to do these, the software will automatically define them assuming linearity (resulting in equally spaced boundary values) between the reference value and the outer boundary levels.

Is it possible to extract the values from the uncertainty results (the uncertainty visualized as bars)?

Reply: This is not possible in the current version. In the forthcoming version (version 1.3) this will be a possibility.

Where did you get the information for the OSPAR indicators (e.g., MTI)?

Reply: For the MTI, the metadata used in NEAT originated from different sources such as internal OSPAR meeting documentation from ICG-COBAM 2013, provided as background information to DEVOTES , Indiseas project information and from the literature (e.g. Rossberg et al., 2011; Pauly et al., 1998).

Note: The OSPAR process progressed and the internal documents were updated and the indicator has evolved since. An update should be proposed by the COBAM indicator leads and the information should be sent back to NEAT developers.

Can we (from EcApRHA) propose modifications for upcoming NEAT versions?

Reply: Yes, modifications and feedback on the software is welcome and is important for the development of the tool.

Note: Modifications can be done directly by the user in their own software. However, in order for the information to be integrated into the master database, this needs to be communicated to the DEVOTES group.

Methodological-Ecological questions:

How to define regions for the PP indicator (a general discussion about this would be interesting)? PP is expected to be highly regional. For example, coastal areas are typically exposed to larger nutrient loadings compared to offshore areas, and will therefore show a higher PP rate. Similarly, upwelling regions will naturally have larger nutrient levels and higher PP. One problem to resolve is therefore how to define these regions. For this purpose, eco-hydrodynamical water masses has been suggested to be possible definable regions (Fig. 11).

Does it exist previous plankton indicators in NEAT, and how did you deal with defining the spatial assessment units?

Reply: This was a more general discussion point where Torsten Berg shared his point of view.

From a scientific point of view, Torsten agreed on defining the spatial assessment units based on the eco-hydrodynamic regions (as we suggested below). However, the problem is that other indicators may not have the same assessment regions, and if the SAU are not the same for the different indicators the assessment loses its value. For example, if the user would allow individual SAU for each indicator, the final assessment would consist of many small areas with a mix of indicators, which would result in a useless result.

Instead, it is important to define common assessment units that can be applied for all indicators. Until now, these defined areas are typically done from the management organization, and this is probably also what OSPAR will do. For example, HELCOM has predefined assessment regions within the Baltic Sea. Exactly how each indicator was calculated within these regions, Torsten does not know. He suspects, however, that some type of average value was used.

New questions to be addressed to NEAT developers

Question: Is it / will it be possible to assign weights to different indicators based on how responsive they are to a specific pressure before aggregating them so that the GES assessment decision is more based on the result of that specific indicator?

7 Glossary

In the following section, we define the main terms that are used in this document to Integrate ecosystem components. The following table synthesises the various definition sources used:

Terms defined	Source of the definition
Environmental Status	Directive 2008/56/EC
Good Environmental Status (GES)	Directive 2008/56/EC
NEAT	*www.devotes-project.eu/neat-manual-v1-2*
Ecosystem Components	*Patricio et al. 2014 ; www.devotes-project.eu/neat-manual-v1-2*
Habitat	*www.devotes-project.eu/neat-manual-v1-2*
Indicators	*www.devotes-project.eu/neat-manual-v1-2*
Nested Structure	*www.devotes-project.eu/neat-manual-v1-2*
Reference Point/ condition	Directive 2008/56/EC *andBorja et al. 2012*

Environmental Status means the overall state of the environment in marine waters, taking into account the structure, function and processes of the constituent marine ecosystems together with natural physiographic, geographic, biological, geological and climatic factors, as well as physical, acoustic and chemical conditions, including those resulting from human activities inside or outside the area concerned;

Good Environmental Status (GES) means the environmental status of marine waters where these provide ecologically diverse and dynamic oceans and seas which are clean, healthy and productive within their intrinsic conditions, and the use of the marine environment is at a level that is sustainable, thus safeguarding the potential for uses and activities by current and future generations;

NEAT and associated words definition

NEAT: The NEAT (Nested Environmental status Assessment Tool) is an integrative tool for the assessment of the environmental status of an ecosystem. This software tool is a flexible application implementing the biodiversity assessment tools developed in work package6 (Integrative assessment of biodiversity) of the DEVOTES project (http://www.devotes-project.eu).

Ecosystem components: The term ecosystem component (mentioned in Annex VI of the Directive 2008/56/EC: 'measures that influence the degree of perturbation of an ecosystem component' and 'tools which guide human activities to restore damaged components of marine ecosystems') includes both biota and habitats as parts of the ecosystem. The habitats include the non-living or physico-chemical factors and the basic compounds and elements of the environment. The biotaconsists of the living parts of the environment, including the association of a lot of interrelated populations that belong to different species inhabiting a common environment. *In the NEAT tool*, different ecosystem components have been defined such as birds, fish, benthic vegetation or pelagic organisms. However, additional components can be added throughout the assessment process when using the NEAT tool.

Habitat: The term 'habitat' has several meanings in common usage linked to the biotic and abiotic environment. The use of the term 'habitat' in the EcApRHA project, taking into account food webs, pelagic and benthic habitat work package indicators developed, refers to the environment a species or community of species inhabit/occupy at a particular stage in its life cycle. *In the NEAT tool*, the habitat (e.g. pelagic, reef) is hierarchically defined under a spatial assessment unit (SAU). For each SAUs, one or more habitats can be assigned. These habitats are themselves nested and hierarchically structured so an indicator can be assigned to one individual habitat or to more than one habitat.

Indicators: Are distinct features that help quantify descriptors outlined within the MSFD. *In the NEAT tool*, the current version integrates an indicator catalogue (Teixeira et al. 2014) as a source for choosing predefined indicators for the biodiversity assessment. However, additional indicators can be added throughout the assessment process when using the NEAT tool.

Nested structure. The principle applied in the NEAT tool is a hierarchical, nested structure of SAUs and habitats (Fig. i). In later releases, this will be expanded to ecosystem components. The order of these hierarchies is such that the assessment begins with the hierarchically nested SAUs.

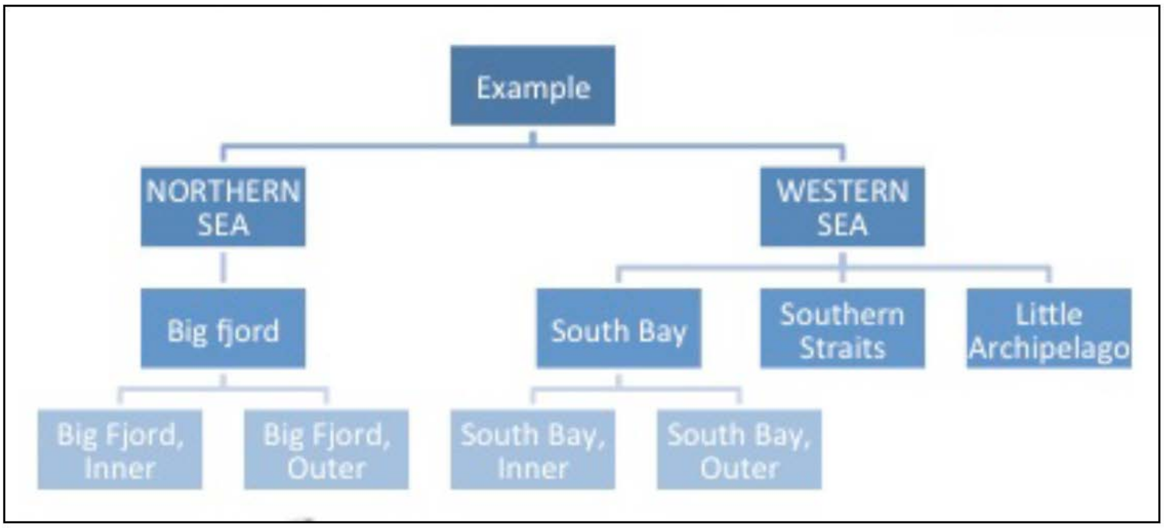


Figure i: Example of spatial assessment units (SAUs) that are nested and hierarchically structured. The SAUs on the higher level always comprise all the units below in the hierarchy. Figure extracted from the NEAT manual available on: http://www.devotes-project.eu/neat-manual-v1-2/

Reference point (also called reference value or reference level): a reference point is the condition of the indicator in the absence (or reduced presence) of human pressure or in pristine areas. Reference conditions are optimally defined/described from the following combined data:

Best acquired from multiple sites with similar physical characteristics, within an ecoregion and habitat type;

That ideally represent minimally impaired or undisturbed conditions (i.e. absence or minimal human pressure);

that provide an estimate of the variability in biological communities and habitat quality due to natural pysical and climatic factors

For more details on setting reference conditions and environmental targets, see Borja et al. 2012

8 References

Barnes, M.K., Tilstone, G.H., Suggett, D.J., Widdicombe, C.E., Bruun, J., Martinez-Vicente, V., Smyth, T.J. 2015. Temporal variability in total, micro- and nano-phytoplankton primary production at a coastal site in the Western English Channel. *Progress in Oceanography* 137: 470-483.

Bopp, L., Respandy, L., Orr, J.C., Doney, S.C., Dunne, J.P., Gehlen, M., Halloran, P., Heinze, C., Ilyina, T., Séférian, R., Tjiputra, J. and Vichi, M. 2013. Multiple stressors of ocean ecosystems in the 21th centry: projections with CMIP5 models. *Biogeosciences* 10: 6225-6245.

Borja, Á., Dauer,D. M., Grémare, A. 2012. The importance of setting targets and reference conditions in assessing marine ecosystem quality. *Ecological Indicators* 12(1):1-7.

Breed, G.A., Jackson, G.A., Richardson, T.L. 2004. Sedimentation, carbon export and food web structure in the Mississippi River plume described by inverse analysis. *Marine Ecology Progress Series* 278:35–51.

Cadee, G. and Hegeman, J. 2002. Phytoplankton in the Marsdiep at the end of the 20th century; 30 years monitoring biomass, primary production, and Phaeocystis blooms. *Journal of Sea Research.* 48: 97–110.

Cole, B. and Cloern, J. 1987. An empirical model for estimating phytoplankton productivity in estuaries. *Marine Ecology Progress Series.* 36: 299–305.

Elliott, S.A.M., Arroyo, N.L., Safi, G., Ostle, C., Aubert, A., Artigas, F., Budria, A., Guérin, L., Rombouts, A.I., McQuatters-Gollop, A., Padegimas, B., Corcoran, E. 2017. "Working towards an ecosystem perspective: Proposals for the integration of pelagic, benthic and food web indicators", as a contribution to the EU co-financed EcApRHA project (Applying an ecosystem approach to (sub) regional habitat assessments), deliverable No. 4.1

EU, 2008. Directive 2008/56/EC of the European Parliament and of the Council of17 June 2008 establishing a framework for community action in the field ofmarine environmental policy (Marine Strategy Framework Directive). *Official Journal of the European Union* L 164/19.

EU, 2010. Decision 2010/477/EU. Commission Decision of 1 September 2010 on criteria and methodological standards on good environmental status of marine waters. *Official Journal of the European Union* L 232/14.

Gailhard-Rocher, I., Huret, M., Lazure, P., Vandermeirsch, F., Gatti, J., Garreau, P., Gohin, F. 2012. Identification de « paysages hydrologiques » dans les eaux marines sous juridiction française (France métropolitaine). *Ifremer report*, 48 pp.

ICES. 2013. Report of the ICES Advisory Committee 2013. *ICES Advice*, book 7.

Joint, I., Wollast, R., Chou, L., Batten, S., Elskens, M., Edwards, E., Hirst, A., Burkill, P., Groom, S., Gibb, S., et al. 2001. Pelagic production at the Celtic Sea shelf break. *Deep-Sea Research II*48: 3049-3081.

Padegimas, B. Artigas, F. Arroyo, N.L. Aubert, A. Budria, A. Capuzzo, E. Corcoran, E. Elliott, S. A. M. González-Irusta, J. M. Guérin, L. Judd, A. Kromkamp, J. McQuatters-Gollop, A. Meakins, B. Niquil, N. Ostle, C. Pesch, R. Preciado, I. Safi, G. Schmitt, P.A. Serrano, R. Thorpe, A. Torriente, Vina-Herbon, C. (2017), *"Action Plan for the further implementation of habitat and food web indicators and progressing integrated assessments in OSPAR (sub) regions"* as a contribution to the EU Co-financed EcApRHA project (Applying an ecosystem approach to (sub) regional habitat assessments), deliverable No. 5.6.

Pauly, D., and Watson, R. 2005. Background and interpretation of the Marine Trophic Index as a measure of biodiversity. *Philosophical Transactions of the Royal Society of London, Series B*, 360: 415–423.

Postma, H. 1954. Hydrography of the Dutch Wadden Sea. *Archives Neerlandaises de Zoologie*10: 405–511Regaudie-de-Gioux, A., Lasternas, S., Agustí, S. and Duarte, C.M. 2014. Comparing marine primary production estimates through different methods and development of conversion equations. *Frontiers in Marine Science*. 1:19.

Safi, G. et al. *In prep*. *Preliminary title*:Describing the Bay of Biscay's continental shelf food-web using an OSPAR common indicator: the Mean Trophic Level indicator.

Smaal, A.C., Schellenkens, T., van Stralen, M.R. and Kromkamp, J.C. 2013. Decrease of the carrying capacity of the Oosterschelde estuary (SW Delta, NL) for bivalve filter feeders due to overgrazing? *Aquaculture*, 404-405: 28-34.

Teixeira, H., Berg, T., Fürhaupter, K., Usitalo, L., Papadopoulou, N., Bizsel, K.C., Cochrane, S., Churilova, T., Heiskanen, A.S., Uyarra, M.C., Zampoukas, N., Borja, A., Akcali, B., Andersen, J.H., Beauchard, O., Berzano, M., Bizsel, N., Bucas, M., Camp, J., Carvalho, S., Flo, E., Garces, E., Herman, P., Katsanevakis, S., Kavcioglu, R., Krause-Jensen, D., Kryvenko, O.,Lynam, C., Mazik, K., Moncheva, S., Neville, S., Ozaydinli, M., Pantazi, M., Patricio, J., Piroddi, C., Queirós, A.M., Ramsvatn, S., Rodriguez, J.G., Rodriguez-Ezpeleta, N., Smith, C., Stefanova, K., Tempera, F., Vassilopoulou,V., Verissimo, H., Yılmaz, E.H., Zaiko, A.,Zenetos, A. 2014. Existing biodiversity, non-indigenous species, food web and seafloor integrity GEnS indicators. *Deliverable D3-1 within the DEVOTES project*.

Tiselius, P., Belgrano, A., Andersson, L., Lindahl, O. 2015. Primary productivity in a coastal ecosystem: a trophic perspective on a long-term time series. *Journal of Plankton Research* 0: 1-11

Van Leeuwen, S., Tett, P., Mills, D., Van der Molen, J. 2015. Stratified and non-stratified areas in the North Sea: Long-term variability and biological and policy implications. *Journal of Geophysical Research: Oceans*. 4670-4686.

Vezina, A.F., Platt, T.C., 1988. Food web dynamics in the ocean. I. Best estimates of flow networks using inverse methods. *Marine Ecology Progress Series*42, 269–287.

ANNEXES

Annex 1

WORK FLOW IN NEAT

We are here presenting the work-flow in NEAT using one of our case studies. This is not complete in any sense, and we therefore recommend the user to turn to the NEAT work manual and to see the NEAT webinar for detailed information (http://www.devotes-project.eu/neat/). In the end of this work-flow section, the results of the assessment can be found (in point 3).

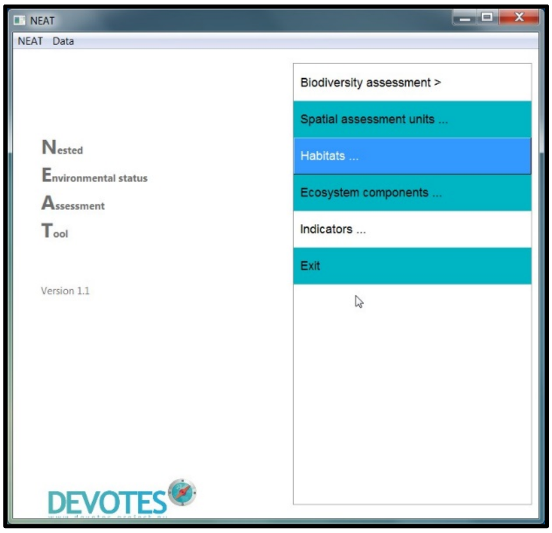


Figure 1: Main window with the menu to create a new assessment.

1) The first step in any assessment is to define any new Spatial Assessment Unit (SAU), habitat and/or ecological group. This is only relevant if the indicators to be assessed is not part of the existing settings of the DEVOTool (this can be explored within each heading in the main window (Fig. 1), i.e., “Spatial assessment units…”, “Habitats…” etc.).

To add a new component, simply press the heading of interest, e.g., “Habitats…”.

In the new Habitat window opened (Fig. 2), press the + sign in the left hand corner to add a new component.

In the new “New Object” window, add the name of the new habitat, e.g., “ecosystem” (Fig. 2), and choose the parent location.

Note: To avoid the software to crash during a final assessment, embed the new components within the DEVOTool headings. This is done by choosing the parent location to be“DEVOTool habitat” (Fig. 2, “New Object” window).

Repeat the same procedure for each potential new component (e.g., SAU, Habitat, Ecosystem component).

Note: A new indicator can be added from the main window. This is not recommended, however, but it is better to do in a later stage of the analyses (this will be explained in detail).

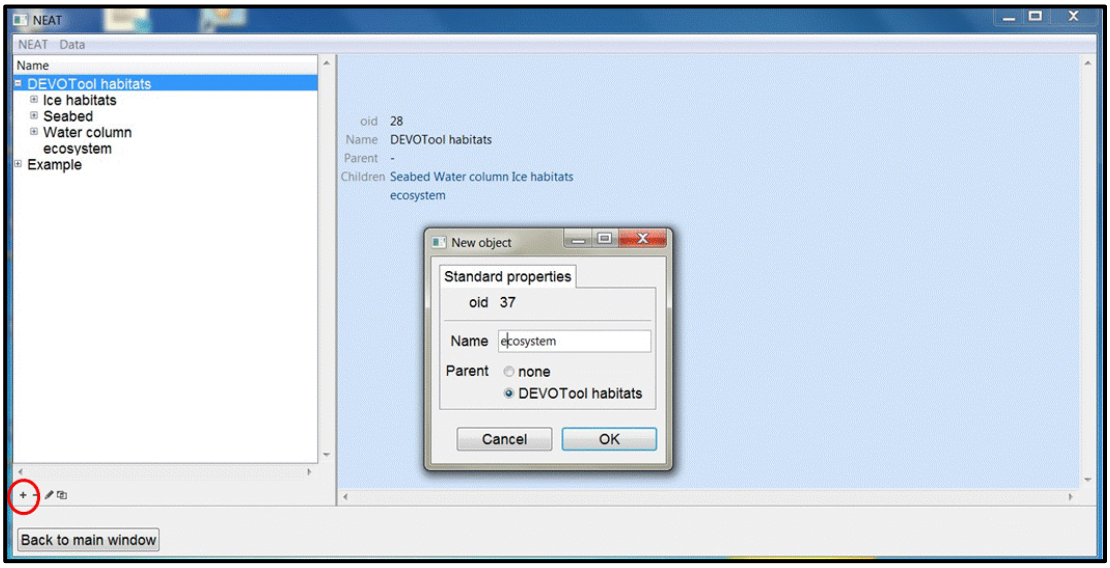


Figure 2: “Habitats” widow, with the “new object” window opened. This is opened by pressing the + sign in the left-hand corner (encircled in red).

In this work-example, we defined:

Habitat: ecosystem (Fig. 2)

Ecosystem component: food web (Fig. 3upper panel)

SAU: Bay of Biscay North and Central (VIIIab). This was located within the “North-East Altantic Ocean” > “Bay of Biscay and the Iberian Coast” (Fig. 3lower panel).

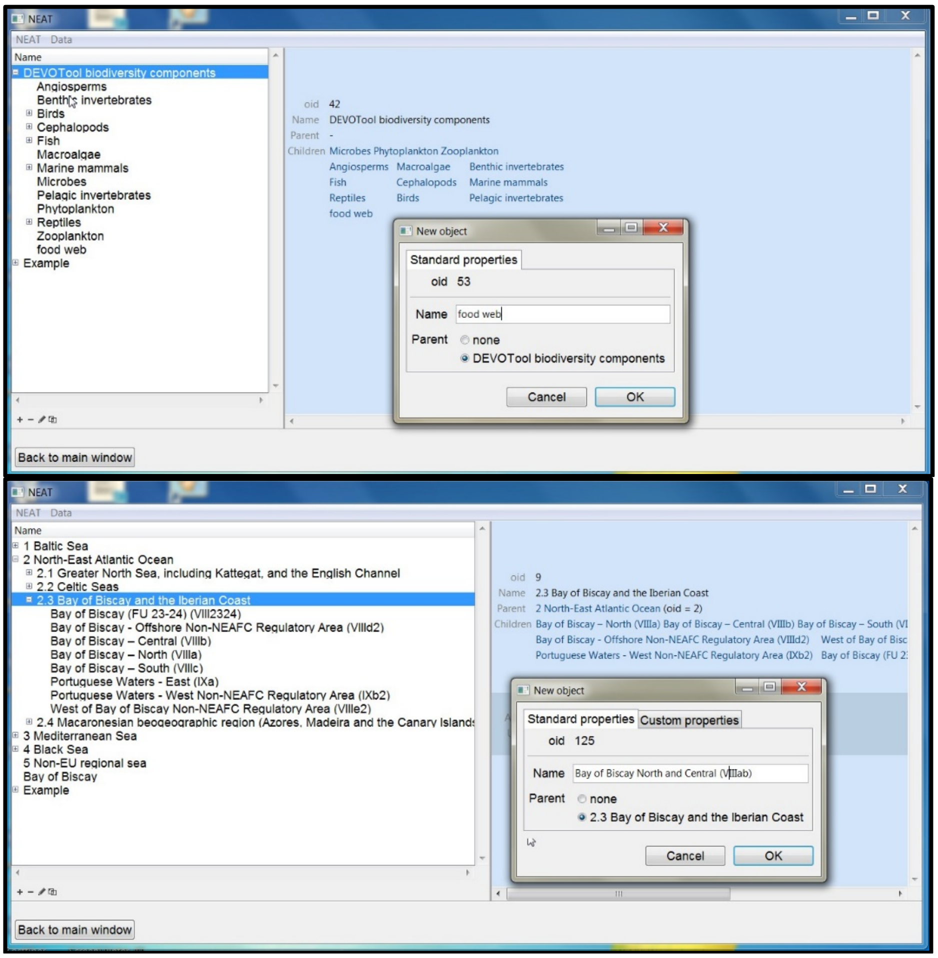


Figure 3: "Ecosystem component" and "Spatial Assessment Unit" widow, with the "new object" window opened. This is opened by pressing the + sign in the left-hand corner.

2) The second step is to open up the assessment window, enter the relevant settings, and run the assessment. This includes defining any new indicator, and set the boundary levels of each indicator.

From the main window, choose "Biodiversity Assessment" (Fig. 1).

In the new "Biodiversity Assessment" window, choose "New Assessment".

In the "New Assessment" window, choose the main spatial assessment unit of interest, e.g., "North-East Atlantic Ocean". This will open up another new window, which is the window for the Assessment settings (Fig. 4). In this window the user will be able to locate their newly defined components, and to enter new indicators.

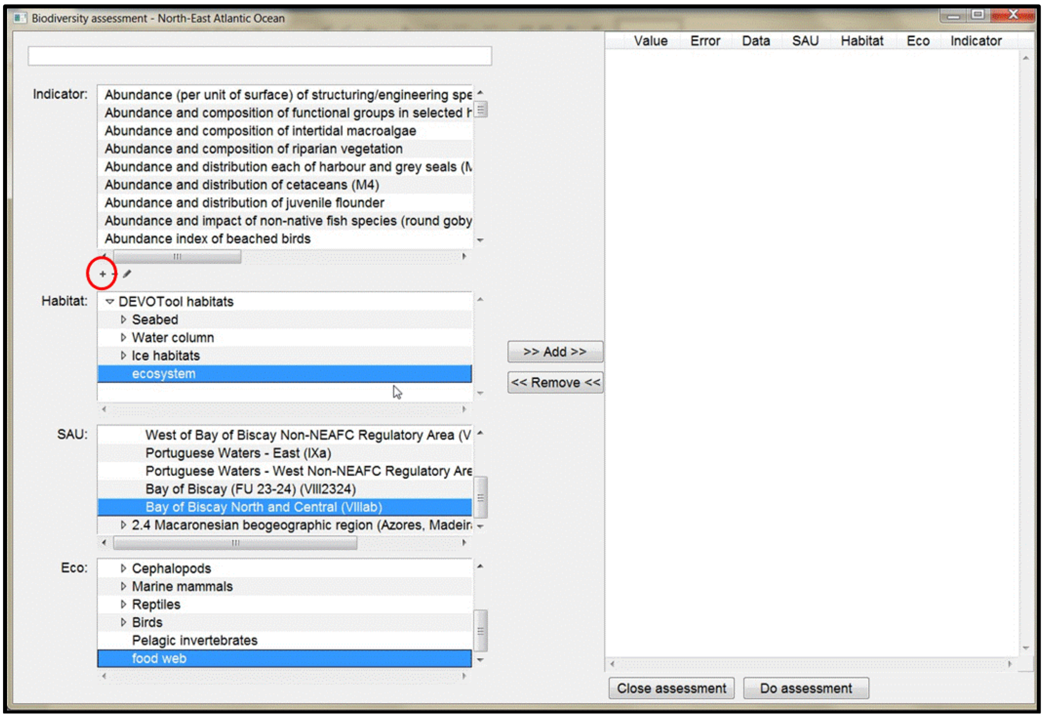


Figure 4: Window for the Assessment settings. The newly defined components (i.e., ecosystem, Bay of Biscay North and Central, food web) are visible in each respective subheading (underlines in blue).

From this window, the new indicator can be added (or any existing indicator can be modified). Add a new indicator by pressing the + sign below the curtain for the "indicator" subheading (Fig. 4, encircled in red). If you want to modify (or set the boundary levels) an existing indicator, press the pencil sign.

A "New indicator" window will open which will take you through several steps:

Step 1 of 4. Enter the name of the indicator, e.g., Mean Trophic Level, and which descriptor it belongs to, in this case the Food webs descriptor. Press Next (Fig. 5, upper panels).

Step 2 of 4. Choose the habitats associated to the indicator. Press Next (Fig. 5, upper panels).

Step 3 of 4. Choose the ecosystem component associated to the indicator. Press Next (Fig. 5, upper panels).

Step 4 of 4. Specify the indicator's status classifications. To specify the classification, press the + sign in the left hand corner (Fig. 4, lower panel, encircled in red).

In the new "Classification" window, locate the SAU, Habitat, and Ecosystem component (Eco) of interest, e.g., Bay of Biscay North and Central (VIIIab), ecosystem, and food web, respectively.

Enter the pre-defined boundary classes. In our case, these are the values given in table 8. Press OK.

Note: Each hierarchical level of an indicator can have different boundary levels (e.g., the North Bay of Biscay can have different boundary levels than the Central Bay of Biscay). This will be done from the "Classification" window.

Repeat the procedure for each spatial level/hierarchy, and for each indicator.

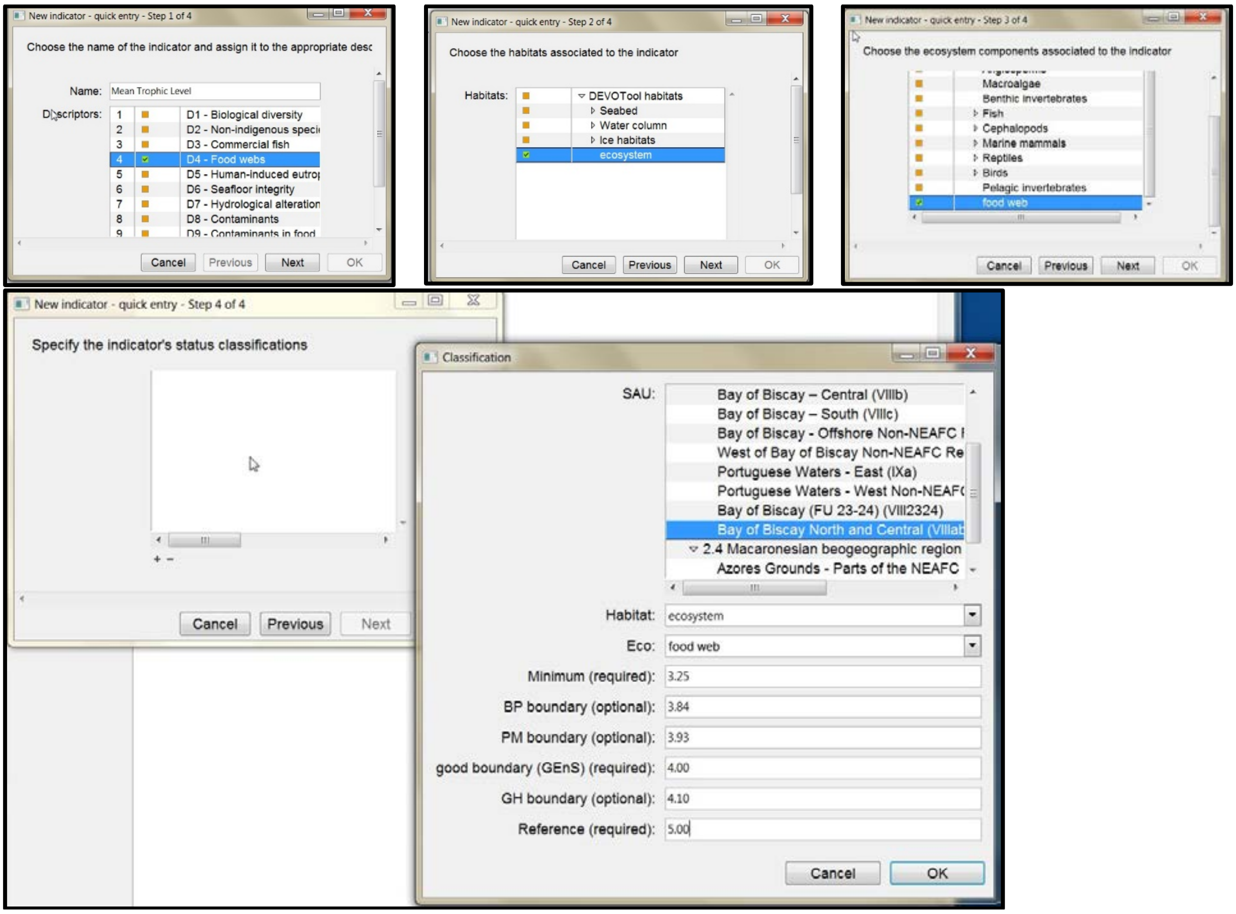


Figure 5: Windows for new indicator (upper panels) and classification (lower panel). The classification window shows the boundary levels entered for the Mean Trophic Level 3.25 (table 8).

After having entered or modified all indicators, it is time to pick out the indicators for each spatial assessment unit in the assessment (Fig. 6, upper panel).

Mark the indicator, Habitat, SAU, and Ecosystem component (Eco) of interest and press >> Add >>.

Do this for each indicator and spatial assessment unit of interest. E.g., in this work-example we run the assessment on the MTL 3.25 and PP for the Bay of Biscay North and central only (Fig. 6, lower panel).

Enter the indicator value and its associated error (as in table 9) in the columns "value" and "error" (encircled in red).

<u>Note</u>: The + sign in front of any of the entries indicates that all basic information needed is entered, hence, if an assessment can be done or not. If for example the boundary classes are missing, a – sign will be visible instead.

When all indicators and spatial assessment units of interest is entered, and the indicator value and its error, run the assessment by pressing Do assessment.

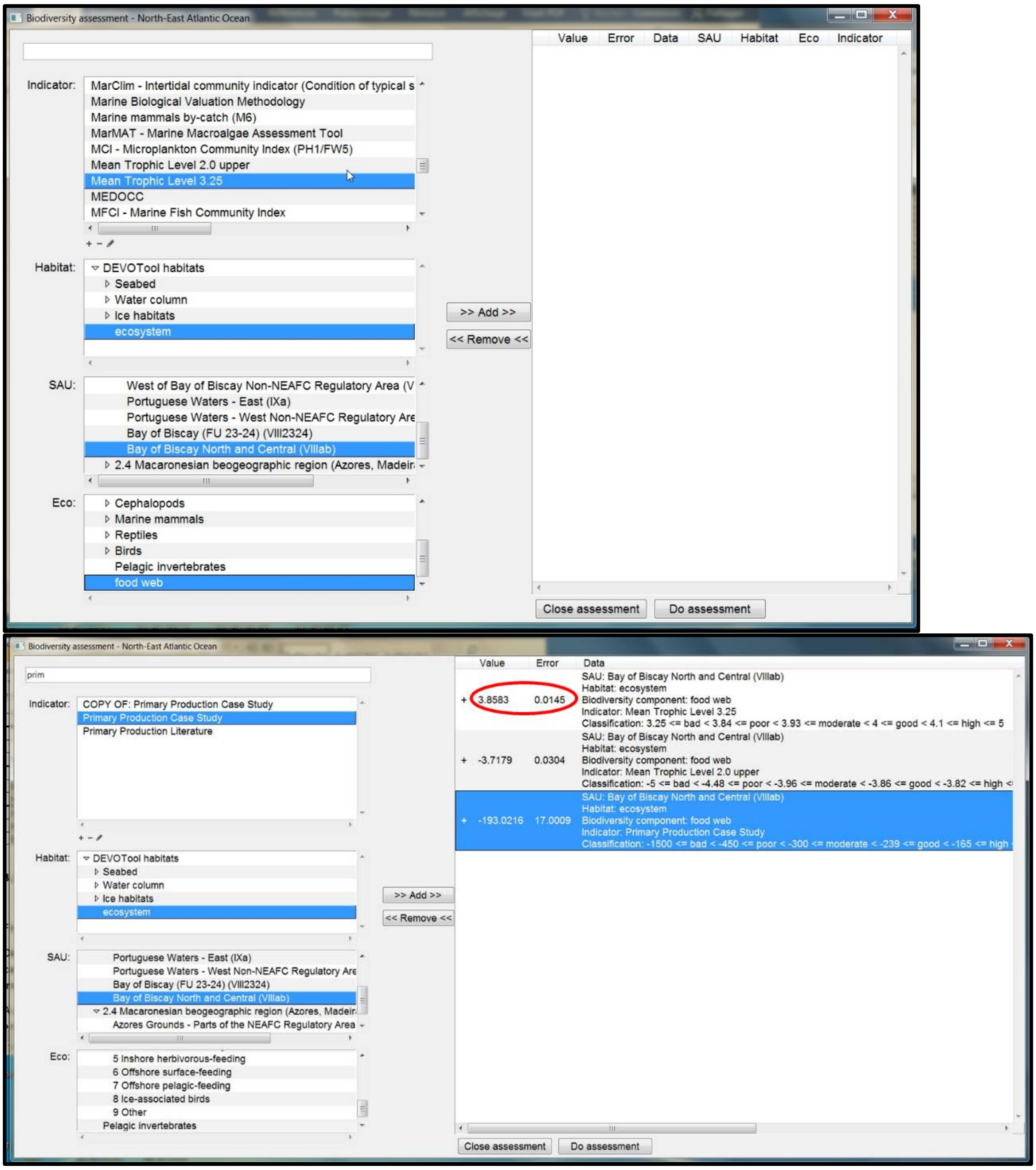


Figure 6: Windows for Biodiversity assessment.

3) The third step is to evaluate the result, and include potential weighting-factors.

In the "result" window, the assessment NEAT value with its associated status class (also indicated by the colour) and uncertainty can be found in the left hand columns (Fig. 7). In an assessment where several ecological components and habitats been analysed, the status of each respective unit can be evaluated. This makes it easier for the user to identify the sensitive areas.

Note: In our example, since we ran the assessment on only one SAU, only one result will be visible. However, if the user defines several spatial levels, a NEAT result will be generated for each spatial level.

Note: When several spatial levels been analysed, weighting can be done by SAU or habitat area. This can be done if e.g. the user believes one assessment area is more or less important than others. For detailed information about the weighting, we refer to the working manual.

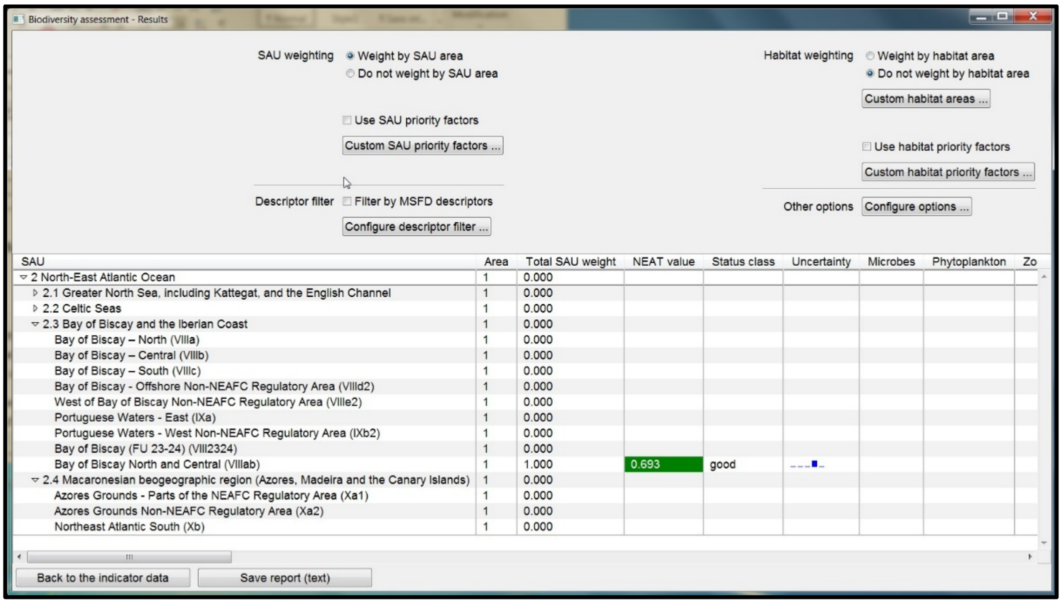


Figure 7: Windows for Biodiversity assessment.

ISBN: 978-1-911458-28-9

Publication Number: EcApRHA 3.5/2017

This report was produced as a result of the EcApRHA (Addressing gaps in biodiversity indicator development for the OSPAR Region from data to ecosystem assessment: Applying an ecosystem approach to (sub) regional habitat assessments) project. The project was co-financed by the European Union (EU). Grant No. 11.0661/2015/712630/SUB/ENVC.2 OSPAR