Indicator-Based Assessment of Marine Biological Diversity–Lessons from 10 Case Studies across the European Seas

Laura Uusitalo, Hugues Blanchet, Jesper H. Andersen, Olivier Beauchard, Torsten Berg, Silvia Bianchelli, Annalucia Cantafaro, Jacob Carstensen, Laura Carugati, Sabine Cochrane, Roberto Danovaro, Anna-Stiina Heiskanen, Ville Karinen, Snejana Moncheva, Ciaran Murray, João M. Neto, Henrik Nygård, Maria Pantazi, Nadia Papadopoulou, Nomiki Simboura, Greta Srebaliene, Maria C. Uyarra and Angel Borja

The Marine Strategy Framework Directive requires the environmental status of European marine waters to be assessed using biodiversity as 1 out of 11 descriptors, but the complexity of marine biodiversity and its large span across latitudinal and salinity gradients have been a challenge to the scientific community aiming to produce approaches for integrating information from a broad range of indicators. The Nested Environmental status Assessment Tool (NEAT), developed for the integrated assessment of the status of marine waters, was applied to 10 marine ecosystems to test its applicability and compare biodiversity assessments across the four European regional seas. We evaluate the assessment results as well as the assessment designs of the 10 cases, and how the assessment design, particularly the choices made regarding the area and indicator selection, affected the results. The results show that only 2 out of the 10 case study areas show more than 50% probability of being in good status in respect of biodiversity. No strong pattern among the ecosystem components across the case study areas could be detected, but marine mammals, birds, and benthic vegetation indicators tended to indicate poor status while zooplankton indicators indicated good status when included into the assessment. The analysis shows that the assessment design, including the selection of indicators, their target values, geographical resolution...
INTRODUCTION

Biological diversity is widely recognized as one of the cornerstones of healthy ecosystems (e.g., Worm et al., 2006). Diversity may safeguard ecosystems against undesired regime shifts (Folke et al., 2004) and guarantee the continued delivery of ecosystem goods and services (Duarte, 2000; Beaumont et al., 2007). The need to maintain biodiversity is also recognized by international legislation (e.g., Convention of Biological Diversity; UNEP, 1992); to European Union (EU) level, the Marine Strategy Framework Directive (MSFD; European Union, 2008) requires its member states to assess the status of marine biodiversity and take action to guarantee that it remains at, or is restored to, Good Environmental Status (GES). A definition of what can be interpreted as good status can be consulted in Borja et al. (2013).

In order to conduct an assessment of status, and to determine the effectiveness of any implemented remedial measures, we need a clear definition of biodiversity and a unified approach for its assessment. In the marine assessments like MSFD, biodiversity is defined on the level of species, communities, habitats, and ecosystems, as well as in the genetic level (Cochrane et al., 2010). Indicators that show the ecosystem response to human pressures form the basis of the tool kit with which we can describe environmental status (Borja et al., 2016). Based on qualitative environmental objectives, targets are set for each indicator which allow policy makers to implement management measures should these not be reached (Borja et al., 2012).

One of the challenges faced during the first round of MSFD initial assessments is the diverging data availability for biodiversity across highly variable systems, but yet an overarching need to conduct compatible assessments across European regional seas (Hummel et al., 2015). European marine ecosystems comprise a complexity and variability both in space and time, ranging from fully saline systems such as in Mediterranean and Atlantic waters to the brackish Baltic Sea, and exposed open water systems such as in the northern Norwegian and Barents seas to fully enclosed systems such as the Black Sea. The levels of available knowledge and data within these systems vary, as well as the biological parameters and indicators used for assessments (Hummel et al., 2015).

The conclusions of the European Commission, in their evaluation of the EU member states’ reports on the initial assessment carried out in 2010–2012 was that there is an apparent lack of coherence and comparability in the indicators used and in the final evaluation of the overall status, between the countries and within all regional seas (Palialexis et al., 2014). Therefore, there is an urgent need for coherent frameworks and methodologies to allow consistent approach in biodiversity assessment across the European Regional Seas. This would also be needed in order to allow coherence in the biodiversity assessments for the EU Birds and Habitats directives and the EU Biodiversity Strategy 2020.

While we could argue that we cannot compare studies if we do not have directly comparable datasets, in practice this is rarely possible, and certainly not at large spatial scales, or involving multiple research institutes and member states. Since there is no single way of describing biodiversity that fits all purposes, and since regional seas have intrinsic differences, we need a pragmatic selection of indicators which are appropriate to the specific questions asked, as well as a flexible and transparent indicator-based tool for assessment of biodiversity status. There is a large number of operational indicators, which have been used to describe the status in different types of aquatic systems (Birk et al., 2012; Borja et al., 2016). As biological diversity is multifaceted, including different taxonomic and functional groups, it cannot be expressed with a single indicator. Consequently, sets of different indicators are needed to cover the broad aspects of biological diversity and it is their combination into a single assessment that becomes a challenge (Borja et al., 2014; Probst and Lynam, 2016). In order to obtain a single overall assessment value, or conclusion, the results of the multiple indicators used in the assessment need to be aggregated, depending on the purpose of the assessment; e.g., if the aim is to inform different stakeholders and to set overall targets for the improvement of the marine environment, or depending on the assessment scale (Borja et al., 2014). Clear and transparent aggregation and integration rules are needed to interpret indicator information onto an environmental status assessment (see Borja et al., 2014 for a review on integration methods).

A variety of assessment tools enabling the integration of indicators already exists (see e.g., HELCOM, 2009a; Andersen et al., 2014; Borja et al., 2016). However, only few of them have treated biological diversity in a comprehensive way, have been tested broadly (i.e., outside the region in which they have been developed), or consider the complexity at an adequate level of detail for the spatial scale for which they are applied. To overcome these issues, in the context of the EU funded project DEVOTES (DEVelopment Of innovative Tools for understanding marine biodiversity and assessing GES), the Nested Environmental status Assessment Tool (NEAT; Berg et al., 2016; Borja et al., 2016) has been developed to assess biodiversity status of marine waters under the MSFD. NEAT uses a combination of high-level integration of habitats and spatial units, and averaging approach (Borja et al., 2014), allowing for specification on structural and spatial levels, applicable to any geographical scale.

Keywords: biodiversity, assessment tool, MSFD, environmental status, spatial aggregation, integration, indicator sensitivity
In this contribution NEAT is applied to the assessment of marine biological diversity in 10 different case studies distributed across the European regional seas (Figure 1). The assessment results are discussed, but the main focus of the paper is on: (i) analyzing the outcome of these assessments in light of the practical choices that have to be made to apply this tool, and (ii) proposing best practices for marine biological diversity assessment using this tool.

**MATERIALS AND METHODS**

**Case Study Areas**

The case study areas were selected to represent a wide range of marine systems (Figure 1), with different climatic and hydrographic characteristics as well as exposure to different human activities and management challenges (Table 1). These areas represent a wide range of marine biogeographical areas from subtropical waters to temperate and Arctic, covering the four European regional seas (i.e., Mediterranean, Atlantic, Black, and Baltic Seas). The surface areas of these case studies varied from <3000 km² in Saronikos Gulf (Greece) to >820,000 km² in the Barents Sea (Norway; Table 1). Detailed descriptions of the case study areas, with relevant references, can be found in Supplementary Material (S1–S10).

**NEAT**

NEAT is a structured, hierarchical tool for making marine status assessments (Berg et al., 2016; Borja et al., 2016), and freely available at www.devotes-project.eu/neat. In NEAT, the study area can be divided into hierarchical spatial assessment units...
| Case study name                      | Code | Regional Sea, subregion     | Salinity range | Sea Surface temperature range (°C) | Mean/max depth (m) | Special features                                                                 | Main pressures                                                                 | Biodiversity components in poorest status | Biodiversity components in best status | Observed gaps in biodiversity components | Observed gaps in habitats | Key references                                      |
|-------------------------------------|------|-----------------------------|----------------|-----------------------------------|-------------------|----------------------------------------------------------------------------------|---------------------------------------------------------------------------------|---------------------------------------------|---------------------------------------------|----------------------------------|------------------------------------------|--------------------------------|
| Norwegian Barents Sea–Lofoten       | BSL  | North-East Atlantic, Barents Sea | 34.4–35        | 1.5–7.5, permanent/seasonal ice cover/ice-free areas | 230/500           | Very large ecosystem; polar region                                                | Climate change, fishing                                                           | Benthic vegetation, birds                  | Phytoplankton, Benthic fauna, zooplankton, fish | Lack of indicators for microbes, reptiles, large benthic organisms, cephalopods, alien species | Rocky and mixed substratum and biogenic reef (ex: Lophelia reefs) | (Wassmann et al., 2006)              |
| Gulf of Finland                     | FIN  | Baltic Sea                  | 0–7            | 0–20, seasonal ice cover          | 37/100            | Estuary; Heavy shipping traffic incl. oil tankers; Large drainage basin; Important resting and breeding area for migratory birds | Nutrient inputs, hazardous substances                                            | Birds, water column habitat, benthic fauna, mammals                          | Zooplankton                                 | Hard bottom fauna, microbes, alien species, habitat-level indicators | Hard bottoms                          | Qaveer et al., 2010                     |
| Lithuanian marine waters            | LIT  | Baltic Sea                  | 5–18           | 9.6–10.4                          | 50/125            | High wave exposure                                                               | Tourism, recreation, ports, shipping, urban, industry, agriculture, fisheries    | Phytoplankton, birds, benthic vegetation                                               | Benthic habitat, benthic fauna, benthic vegetation, zooplankton     | Mammals                                 |                                          |                                           |
| Kattegat                            | KAT  | Baltic Sea, North Sea       | 18–32 (surface)| 2–19 (surface)                    | 23.9/80           | Transition zone between the Baltic Sea and the North Sea.                        | Nutrient inputs, hazardous substances, fishing, physical modification              | Birds, phytoplankton                                                                | Zooplankton, fish                          | Non-commercial fish, alien species, microbes | Hard bottoms                          | Naturstyrelsen, 2012; Andersen et al., 2013 |
|                                     |      |                             | 25–35 (bottom) |                                   |                   |                                                                                    |                                                                                  | (Continued)                                  | (Continued)                                 | (Continued)                              | (Continued)                             | (Continued)                            | (Continued)                            |
| Case study name | Code | Regional Sea, subregion | Salinity range | Sea Surface temperature range (°C) | Mean/max depth (m) | Special features | Main pressures | Biodiversity components in poorest status | Biodiversity components in best status | Observed gaps in biodiversity components | Observed gaps in habitats | Key references |
|-----------------|------|-------------------------|----------------|----------------------------------|-------------------|------------------|---------------|----------------------------------------|----------------------------------------|----------------------------------------|--------------------------------|-----------------|
| Dutch North Sea | DNS  | North Sea               | 27.6–34.8      | 4–18                             | 6/54              | Shallow and highly productive area, soft bottom sediments, contrasted benthic habitats (stressed in the southern part, rarely disturbed in the northern one) | Beam-trawl fishing, organic loading, harbor infrastructures | Fish | Birds, mammals, benthic fauna | Benthic vegetation | Herman et al., 2014 |
| Basque Coast    | BAS  | North-East Atlantic     | 32.5–35.5      | Range 0–5000                     | High wave exposure, long fetch | Fishing/shipping | Mammals       | Phytoplankton, zooplankton, benthic habitat, water column habitat | Microbes | Deep and abyssal habitats | Borja and Collins, 2004 |
| Portuguese continental subdivision | POR | North-East Atlantic     | 35–36          | Range 0–6000                     | High wave exposure and seasonal upwelling at north-west coast, smoother conditions at south and south-west coast | Coastal trawling (also dredging, sediment disposal, harbors) | Fish, birds, pelagic fauna | Zooplankton | Reptiles, mammals, microbes. Species distribution and coverage area. The current status of all stocks has been assessed, although using secondary indicators | Deep and abyssal habitats and open-sea areas (>20 nm) | MAMAOT, 2012 [https://dl.dropboxusercontent.com/u/103729442/EstrategiaMarinha_subdv_Continente.pdf](https://dl.dropboxusercontent.com/u/103729442/EstrategiaMarinha_subdv_Continente.pdf) | Dupont et al., 2014 |
| Black Sea coast (Large Varna Bay) | BSC  | Black Sea               | 11–18          | 0–28                             | 30                | Hydrologically dynamic area | Tourism industry, urbanization, port activity, aquaculture | Fish, phytoplankton, zooplankton | Water column habitat, benthic habitat | Microbes, alien species, food-web indicators; insufficient fish monitoring | No monitoring in the open sea areas | National Initial Assessment GES Report, 2013 |
| Case study name | Code | Regional Sea, subregion | Salinity range | Mean/max depth (m) | Special features | Main pressures | Biodiversity components in poorest status | Biodiversity components in best status | Observed gaps in biodiversity components | Observed gaps in habitats | Key references |
|-----------------|------|-------------------------|---------------|-------------------|------------------|----------------|------------------------------------------|--------------------------------------|----------------------------------------|----------------------------|----------------|
| Saronikos Gulf  | SAR  | Mediterranean (Eastern Mediterranean/ Aegean Sea) | 38–39 12–28   | 100/450           | It is the natural marine gateway of the city of Athens and Piraeus harbor. It receives the effluents of the central sewage outfall of Athens through a deep underwater outlet. | Activities: Waste Water Treatment Plant (WWTP), Shipping, tourism, fishing. Pressures: input of organics/nutrients, contaminants, habitat loss, resource exploitation, alien species, litter | Mammals, fish | Reptiles | Birds, zooplankton | | Simbourea et al., 2014, 2015; SoHelME, 2005; SoHelFI, 2007 |
| Adriatic Sea    | ADR  | Mediterranean | 37–39 18–25 | 35/1200 | Cyclonic circulation; input from the Po river | Organic loading; overexploitation; bottom trawling | Mammals, phytoplankton | Fish, birds, zooplankton | Reef and mud habitats | | Artegiani et al., 1997a; Artegiani et al., 1997b; UNEP/DEP/MED WG.408/Inf.14, 2013 |
(SAU) and habitat types (HBT); e.g., SAU “archipelago zone” could include “inner archipelago” and “outer archipelago” as lower-level SAUs, and they, in turn, could include, e.g., water bodies as yet lower-level SAUs. Similarly, the HBT “seafloor” could include HBTs “soft bottom” and “hard bottom,” which again could be further sub-divided (Figure 2). NEAT classifies the status of each SAU based on indicators that have been defined for that SAU; if one SAU has indicators describing different HBTs, the status of each HBT within a SAU is assessed first, and each HBT is then given equal weight in assessing the status of the SAU. The overall assessment is an average of the SAUs, weighted by their surface areas (km$^2$). Other weighting schemes can be applied, if desired.

Each indicator must be explicitly linked to a SAU and a HBT—the same indicator, e.g., “the maximum depth of seaweed,” can be included multiple times for multiple SAUs and HBTs if it has been assessed for multiple areas. These instances of indicators are called “indicator values” in this paper, while the indicators describing a certain ecological concept, e.g., the growth depth of a macrophyte species, or the reproduction rate of a bird species, are called “unique indicators.”

In order to aggregate indicators by weighted average, it is necessary to transform all indicators to a common scale. In NEAT, indicators are transformed into values that range from 0 to 1 using a continuous piecewise linear function. On this scale, the value of 0.6 corresponds to the boundary between good (>0.6) and not good (<0.6) status. Transformation to this scale is defined by specifying the values of the indicator in the original measurement scale, which corresponds to the transformed values of 0, 0.2, 0.4, 0.6, 0.8, and 1.0. Though the transformation function is piecewise linear, the definition of 5 segments allows a reasonable approximation to non-linear functions. These five segments are also used here for illustrative purposes, and they are called bad/poor/moderate/good/high classes, although it is recognized that the boundary between GES and non-GES lies between the “moderate” and “good” classes.

**Indicator Selection and Specification**

The indicators used for this assessment represent the best available data and expertise for the six biological descriptors of the MSFD [i.e., D1 (biodiversity), D2 (non-indigenous species), D3 (commercially important species), D4 (food webs), D5 (eutrophication), and D6 (sea floor integrity)] in each case study area. These indicators include the national and regional indicators used for the MSFD assessment, and indicators derived from scientific literature and expertise. They have been selected to be representative of various biodiversity components, habitats, and geographical areas relevant for each case study area; however it is possible that no indicators exist to be used for some relevant components. The list of indicators included in each case study is available in Supplementary Material S11.

![Figure 2](image-url)
Each indicator is associated to an ecosystem component class that describes the ecosystem component that the indicator describes. In this study, 12 ecosystem components were defined in order to accommodate all indicators used in all of the case studies. These components were phytoplankton, zooplankton, fish, reptiles, marine mammals, birds, benthic fauna, benthic vegetation, pelagic fauna (composite indicators consisting of data from multiple pelagic fauna groups), all taxa (composite indicators consisting of data from multiple taxa), benthic habitat, and water column habitat. The latter two components gathered indicators related to physico-chemical conditions of the habitat, necessary to maintain life (e.g., oxygen or nutrients), whilst the “all taxa,” benthic fauna, and pelagic fauna groups included composite indicators encompassing many species groups; the other nine ecosystem components were taxonomic groups.

Biodiversity Status
The status of the biological diversity was assessed for each case study area using NEAT. The analysis provides an overall assessment for each case study area and a separate assessment for each of the ecosystem components included in the assessment. The final value has an associated uncertainty value, which is the probability of being in a determinate class status (GES/non-GES). This uncertainty was determined by the standard error linked to the indicator values (Carstensen and Lindegarth, 2016).

Evaluation of Assessment Design and Its Effects on the Status Assessment
The application of NEAT to a broad range of marine regions provides an opportunity to test and compare the NEAT assessment approaches and evaluate the consequences of design choice for the general environmental status assessment. How the available data are combined within the tool might have consequences on the results of the status assessment of biodiversity (Borja et al., 2014; Probst and Lynam, 2016). Therefore, one of our aims is to evaluate the consequences of the way the assessment was designed on the general assessment result.

NEAT gives a framework to organize the assessment, but it does not prescribe the number of assessment components, i.e., indicators, SAUs, HBTs, or ecosystem components to be used in an assessment. The user has the option to organize the different components of NEAT depending on the case, e.g., the morphological characteristics of the area, availability and resolution of data, and how the selected local indicators are defined.

In order to describe the assessment design, the following key components were summarized for each case study: (i) the total number of SAUs and how many hierarchical SAU levels there are, (ii) the total number of HBTs and their hierarchical levels, (iii) the number of ecosystem components covered by the indicators, (iv) the number of unique indicators (i.e., not repetition of the same indicator on a different spatial unit), as well as (v) the quantity of data, defined as the number of different indicator values (e.g., if the same indicator is defined separately for five different SAUs, they would comprise five indicator values).

NEAT assigns weights to the indicators based on the SAU and HBT that they represent (see Section Evaluation of the Assessment Results). The SAUs are weighted according to their surface area and the HBTs are weighed equally within a SAU. Therefore, the indicator values contribute to the assessment with different weights, the highest weight being assigned to an indicator representing a large SAU with a small number of indicators, and within it a HBT with a small number of indicators. The relative weights of the indicator values were used to identify the indicators that contribute 90% of the weight of the final assessment. In addition, the relative weight of each ecosystem component in each case study assessment was calculated. These summary statistics highlighted differences in aggregating information among case studies.

To test the sensitivity of the case study assessments to the selection and number of indicator values, a sensitivity analysis was performed by running the assessment using randomly selected indicator values. The number of indicator values included into the assessment varied from 1 to the maximum number of indicators in the case study minus one. This process was repeated 100 times for each number of indicator values. For example, take a case study with 120 indicator values. First, one random indicator value is selected and the assessment is done using only that indicator. This procedure is repeated 100 times. Then, two indicator values are picked at random, and the assessment is run using them; this again is repeated 100 times. This procedure is repeated for all numbers of indicator values up to 119. This results in a large number of values whose divergence can be analyzed to see if any patterns can be identified.

RESULTS
Assessment Design
The number of SAUs as well as how many hierarchical levels were used in these varied widely between the case studies. The number of SAUs included in the Gulf of Finland and Portugal continental sub-division cases were much higher (>60) than in all other case studies which included, on average, 9 different SAUs. Excluding these two case studies, larger areas were usually assessed using more SAUs. The number of hierarchical SAU levels varied between 1 and 5, but in 7 out of 10 cases, there were 3 or 4 levels (Table 2, Figure 2). The total number of HBTs included in the assessment varied between 3 and 9, and 9 out of 10 case studies had 2 or 3 hierarchical HBT levels (Table 2). Not all SAUs necessarily included all habitat types, and indicators or data may not exist for all defined HBT types for each SAU. The number of SAU-HBT combinations that were assessed by at least one indicator value, varied between 6 and 132 (Table 2).

The number of ecosystem components included in the analyses varied between 5 and 9, with an average of 7.3 (Table 2). It has to be noted that all ecosystem components identified in this study were not applicable to all areas; an example being reptiles that do not occur in most of the study sites.

The number of unique indicators applied in each case study area varied between 11 and 116 (Table 2, Supplementary Material S11). The number of indicator values varied greatly with 466
values at the higher end in Portugal continental sub-division and between 20 and 200 values in all other case studies (Table 2).

### Biological Diversity Status

The summary of the test NEAT assessments of the 10 case study areas is presented in Figure 3. The assessment resulted in GES for the Basque EEZ and the Barents Sea–Lofoten, with 100 and 66% confidence, respectively, the remaining eight case studies presented non-GES (i.e., bad, poor, or moderate; Figure 3). Lithuanian coast has the potential for being in GES, but with a 66% confidence, respectively, the remaining eight case studies for the Basque EEZ and the Barents Sea–Lofoten, with 100 and 66% confidence of 20% (Figure 3). For the other case studies, this probability of achieving GES was <1% (Figure 3).

The different ecosystem components showed different status in the case study areas (Figure 4). No strong pattern among the ecosystem components could be detected, but some commonalities were found: Indicators based on marine mammals generally indicated degraded situation in 6 cases out of 7 (Figure 4). When included, birds and benthic vegetation indicators as well as water column indicators of physicochemical status also indicated degraded situation in 5 cases out of 7. Indicators encompassing several ecosystem components ("AT," on Figure 4) always indicated degraded situations. On the other hand, indicators of benthic habitats' physico-chemical status and of zooplankton community status indicated GES when they were included in the assessment (Table 1, Figure 4).

### Relative Contribution of Indicator Values and Biodiversity Components

The indicator values contributed differently to the final assessment result (Figure 5); indicator values defined for larger SAUs tend to have more weight, particularly if there are only few indicators defined for these SAUs. In 7 out of the 10 case studies, <10 indicator values already contributed to more than 50% of the final assessment result. For 9 case studies, <50 indicator values contributed to >90% of the final assessment. This 90% of the final assessment was reached with <20 indicator values in five case studies (Figure 5). The five indicator values that made the highest contribution to the final assessments of each case study are listed in Table 3. These indicator values were dominated by mammal, bird, fish, and benthic fauna indicators.

The 12 different ecosystem components’ contribution to the final assessment result did not correspond to the number of indicator values defined for each component (Table 4). For example, most case studies had a large proportion of benthic fauna indicator values (average: 22.4% of indicators values), which ultimately did not reflect proportionally in the final assessment (average contribution: 11.7%). In contrast, the proportion of fish and marine mammals indicator values were lower, but these components contributed to a higher proportion of the final assessment. In five case studies (i.e., Barents-Lofoten, Gulf of Finland, Dutch North Sea, Saronikos Gulf, and Adriatic Sea), “Benthic fauna” was the component with the highest proportion of indicator values (Table 4); the other five case studies each had a different component with the highest number of indicator values. However, in none of the case studies, benthic fauna was the component with highest contribution to the final assessment (Table 4); in five (i.e., Gulf of Finland, Dutch North Sea, Basque coast, Portuguese continental subdivision, and Black Sea coast) and two case studies (i.e., Barents Sea–Lofoten and Adriatic Sea) respectively, fish and mammals were the components carrying the highest weight to the final assessment (Table 4). However, other ecosystem components, that overall did not contribute to many case study assessments, were very relevant for specific case studies (e.g., the composite

---

### Table 2 | Synthesis of the structure used by the different case studies for the nested assessment.

| Case study name | Area (km²) | Number of SAU levels (total number of SAU) | Number of HBT levels (total number of HBT) | Number of SAU*HBT combinations with data | Number of ecosystem components included | Number of unique indicators (see Supplement 11) | Number of indicator values |
|-----------------|-----------|---------------------------------------------|--------------------------------------------|------------------------------------------|----------------------------------------|-----------------------------------------------|---------------------------|
| Norwegian Barents Sea–Lofoten | 821 478 | 4 (13) | 3 (9) | 21 | 7 | 40 | 74 |
| Gulf of Finland | 22 482 | 5 (60) | 2 (3) | 103 | 8 | 25 | 147 |
| Lithuanian marine waters | 6 426 | 2 (4) | 2 (7) | 6 | 9 | 27 | 50 |
| Kattegat | 17 440 | 3 (11) | 2 (7) | 21 | 8 | 31 | 69 |
| Dutch North Sea | 57 000 | 1 (1) | 3 (6) | 6 | 6 | 15 | 31 |
| Basque Coast | 10 794 | 3 (8) | 3 (6) | 22 | 9 | 48 | 109 |
| Portuguese continental subdivision | 268 645 | 4 (61) | 4 (7) | 132 | 7 | 14 | 466 |
| Black Sea coast (Large Varna Bay) | 1 434 | 3 (7) | 2 (4) | 15 | 7 | 35 | 112 |
| Saronikos Gulf | 2 907 | 3 (4) | 3 (6) | 10 | 7 | 17 | 29 |
| Adriatic Sea | 138 600 | 3 (10) | 2 (5) | 17 | 5 | 116 | 177 |
| Mean | 134 721 | 3.1 (17.9) | 2.6 (6.0) | 35.3 | 7.3 | 39.2 | 126.4 |
| Stdv | 255 807 | 1.1 (22.7) | 0.7 (1.7) | 44.3 | 1.3 | 30.7 | 128.8 |

SAU, spatial assessment unit; HBT, habitats. The case studies are ordered according to their latitude.
group “all taxa” in the Saronikos Gulf and benthic habitat in the Lithuanian coast).

Sensitivity Analysis
The sensitivity analysis shows that there are major differences in how much the result varies if only a subset of the indicator values is included in the assessment (Figure 6). For example, if only a small number (close to 0) indicators were included, the assessment results in all studies could be anywhere between high and bad status, except in Barents Sea and Portuguese continental subdivision, where they could range from poor to high status. As more indicator values are added, the range of outcomes narrows down. However, how steeply that happens when indicator values are added varies between the case study areas (Figure 6).

DISCUSSION
The current NEAT-based assessment demonstrates a large-scale marine biodiversity assessment, providing a feasible solution to the apparent problem pointed out by the European Commission, in their evaluation of the EU member states' reports on the MSFD initial assessments carried out in 2010–2012 (Palialexis et al., 2014). This problem was the apparent lack of coherence and comparability in indicators used and in the final evaluation of the overall status between the countries and within all regional seas (Palialexis et al., 2014). Despite the available guidance and Commission Decision (European Union, 2010) on GES descriptors, criteria and indicators, the overall picture in assessments was patchy and non-coherent (European Commission, 2014). The use of NEAT, and its validation in different regional seas and case study areas, is a crucial contribution from the DEVOTES project to provide a harmonized approach and methodology for a coherent and comparable environmental status assessment across the European regional seas. It also shows that although the regional seas have different characteristics and human pressures impacting those (Claudet and Fraschetti, 2010; Micheli et al., 2013a; Andersen et al., 2015), a coherent assessment framework can be employed to evaluate differences in the environmental status and the ecological components that are impacted by different pressures.

The study and the comparison of the case studies brought into light several issues that need attention in order to improve the coherent and comparable “biodiversity status” assessments of the European regional seas. These issues are related to the data and indicator availability, how the assessments are structured, how the integrative assessment should be structured, and how this structure should be taken into account when defining the spatial resolution and indicator selection of the assessments. The current study revealed that while these assessments could be carried out, there are two major problems in achieving the objectives of the MSFD assessments: (i) there are still multiple gaps in the availability and coverage of indicators in the various areas, and (ii) comparability of the status assessments across different regions would benefit from a more unified assessment framework, even if indicators suitable for each area remained different. NEAT provides a general framework that could be accompanied with guidelines for the selection of SAUs, HBTs, and indicators.

Each of the case studies was initially designed with the best available selection of spatial units, habitats, and indicators, adhering to the NEAT methodology but without specific guidelines for the indicator selection, target level setting, etc. This situation resembles the situation where the new users would start using NEAT on their area. For the purposes of this study, the assessments were evaluated and harmonized to some degree, e.g., if the same indicator appeared in multiple case studies, it was ensured that it was associated to the same biodiversity component (e.g., chlorophyll a levels would be assigned to phytoplankton). Despite this harmonizing, there were major differences in how the case studies were constructed in terms of spatial resolution, habitats, and indicator definition. The current assessment is based on best available data and evaluation of the experts participating within this exercise, and the biodiversity status results of this study should be considered as indicative, not definitive.

The indicators selected for the assessments are designed or adapted for each area separately, including the geographical and habitat specification and the target level, i.e., which values are considered good and which less than good in any given area and habitat. This means that the “good” status is scaled according to the area: In areas with a naturally low biodiversity, lower...
biodiversity is also considered “good” than in areas with naturally high diversity. This makes the assessment relevant for each area, and the result must be interpreted to be in relation to undisturbed condition of that area rather than in absolute terms of diversity.

According to a categorization of rules or methods for combining or aggregating indicators or criteria within a given descriptor (Prins et al., 2013; Borja et al., 2014), NEAT is classified as a high-level integration method which reduces the risks associated to the “one out, all out” principle of the Water Framework Directive approach (Borja and Rodríguez, 2010) while giving an overall and specific (to descriptors and components) assessment.

According to the relevant guidance document for the MSFD (Prins et al., 2013), the spatial scales are not the same for all indicators within the biodiversity descriptor, where depending on the species or habitat a different spatial scale may be used. It is also recommended to address uncertainties and assess confidence of the classification result (as a secondary assessment). In our study, the NEAT software treats equally all assessment elements assigning equal weights, but gives more weight in cases of larger spatial coverage, with higher data representativeness, in that way.

FIGURE 4 | Probabilities for the five environmental status classes separated by ecosystem component for each of the 10 case studies (for codes see Table 1). Ecosystem components considered are: Phytoplankton (PPL), zooplankton (ZPL), benthic vegetation (BV) benthic fauna (BF), fish, birds, mammals (Mamm.), reptiles (Rep.), pelagic fauna (PF), all taxa (AT), water column habitat (WCH), benthic habitat (BH).
FIGURE 5 | Cumulative contribution to the final assessment (cumulative weight) in relation to the number of indicator values. For each case study, values have been ranked from the most important indicator value (highest weight) to the least important value (lowest weight). The X-axis has been cut; Portuguese case study has a total of 466 indicator values. For case study codes, see Table 1.

incorporating the spatial scales issue and the confidence level into the assessment. This could be the reason for which some ecosystem components (e.g., seabirds, mammals, and fishes) have more weight in the final assessment, since they are normally assessed at large scale spatial areas, which have more weight when aggregating (e.g., Saronikos gulf). However, NEAT also includes the possibility to weight indicators differently.

Implications of the Assessment Design

Most of the case study areas lacked indicators regarding one or several biodiversity components and habitats (Table 1, Figure 4), even those that were deemed important in the area. The lack of indicators stemmed either from lack of monitoring data regarding the area or biological diversity component (e.g., birds, reptiles, pelagic fauna), or from obstacles in the indicator development, including the lack of expert time to develop indicators, or insufficient knowledge about the target levels due to lack of long-term or reference condition data (Hummel et al., 2015). In some cases, more basic ecological research is needed in order to understand the ecological processes well enough to develop indicators. In fact, most of the assessments undertaken until now by member states is more qualitative than quantitative (Hummel et al., 2015), representing a challenge for the assessment.

The habitats and biodiversity components for which no indicators are available potentially affect the final assessment result. It is entirely possible that adding even one indicator that would represent a poorly-represented, large area or habitat, would change the overall assessment for better or for worse. Therefore, in order to make a reliable assessment of the status of the biological diversity, the critical gaps in each assessment case need to be evaluated for their potential to affect the overall result. If such high-leverage gaps exist, the assessment result must be taken with caution.

Different indicator values and spatial assessment units had varying weights in the final assessment result in all of the cases (Table 3, Figure 2). The differences in the indicator value weights stem from the fact that the default NEAT assessment first assesses the result for each SAU, giving equal weight to each HBT with similar hierarchy, and combines these SAUs hierarchically so that each SAU is given weight according to its area. Therefore, if a SAU has a large surface area and only a small number of indicators per one or several of its habitat types, these indicator values end up contributing strongly to the final assessment.

This emphasizes the importance of the balanced nature of the indicator set, and particularly the reliable assessment of indicators that are used to assess the status of large areas, and particularly their habitats with only few indicators (Feary et al., 2014). Therefore, particular attention should be paid to both the observed value, the boundary values between the classes, and the uncertainty estimation of these most influential indicator values.

The fact that the SAUs are weighted according to their surface area in the default mode of NEAT also emphasizes the need for careful consideration of the definition of the SAUs. Ideally, the SAUs should be defined in the manner that an indicator value defined for a SAU can be expected to reasonably represent all of the SAU. On the other hand, if the assessment area is split into several sub-SAUs and only a fraction of them actually has indicator data, their value will be generalized to represent the whole super-area in the hierarchical assessment anyway.
### TABLE 3 | List of the top-five indicator values contributing the most to the overall assessment for each case study.

| Spatial Assessment Unit (SAU) | Habitat | Ecosystem component | Indicator name | % contribution to assessment |
|------------------------------|---------|---------------------|----------------|-----------------------------|
| **BARENTS SEA LOFOTEN**      |         |                     |                |                             |
| Norwegian BARENTS SEA–LOFOTEN| sea-ice | Mammals             | Cumulative contrib. of top-five indicators | 56                          |
| Ice-free zone                | pelagic | Birds               | Harp seal, Pagophilus groenlandicus stock size | 16.6                        |
| Seasonal-ice zone            | sea-ice | Mammals             | Kittiwake, Rissa tridactyla breeding success over last 5 year | 16.4                        |
| Offshore (ice-free zone)     | shelf (muddy sediments) | Benthic Fauna | Proportion of non-threatened pagophiletic mammals | 15                          |
|                             |         |                     | 4 indicators related to macrobenthic fauna: ES100, abundance level, evenness and AMBI | 3.7 (each) |
| **GULF OF FINLAND**          |         |                     | Cumulative contrib. of top-five indicators | 60                          |
| Open sea                     | Benthic | Benthic Fauna       | Average regional diversity | 18.4                        |
| Gulf of Finland              | Pelagic | Fish                | Abundance of salmon spawners and smolt | 12.5                        |
| Gulf of Finland              | Pelagic | Fish                | Herring, Spawning stock biomass | 12.5                        |
| Gulf of Finland              | Marine (general) | Mammals | 3 indicators: Gray seal population growth rate, Gray seal pregnancy rate, Ringed seal population growth rate | 8.3 (each) |
| **LITHUANIAN COAST**         |         |                     | Cumulative contrib. of top-five indicators | 47                          |
| Territorial sea              | Benthic | Benthic habitat     | Extent of the seabed significantly affected by human activities | 31                          |
| Territorial sea              | Marine (general) | Others | Biopollution level index (invasive species) | 4                          |
| Territorial sea              | Marine (general) | Birds | Abundance of wintering populations of seabirds: (1) Red-throated Diver + Black-throated Loon, (2) Great Crested Grebe (Podiceps cristatus), (3) Common Merganser (Mergus merganser), (4) Velvet Scoter (Melanitta fusca), (5) Long-tailed Duck (Clangula hyemalis), and (6) Common Goldeneye (Bucephala clangula) | 4 (each) |
| **KATTEGAT DK**              |         |                     | Cumulative contrib. of top-five indicators | 30                          |
| KATTEGAT, central parts      | Marine (general) | Birds | Fulmar winter abundance (encounter rate) | 6.6                        |
| KATTEGAT, central parts      | Marine (general) | Birds | Kittiwake winter abundance (encounter rate) | 6.6                        |
| KATTEGAT, central parts      | Marine (general) | Birds | Guillemot winter abundance (encounter rate) | 6.6                        |
| KATTEGAT, central parts      | Benthic | Benthic Fauna       | BQI            | 5                          |
| KATTEGAT, central parts      | Benthic | Benthic Fauna       | DKI            | 5                          |
| **DUTCH NORTH SEA**          |         |                     | Cumulative contrib. of top-five indicators | 21                          |
| Dutch EEZ                    | Benthic | Benthic Fauna       | Benthic invertebrates total number of species | 16.7                        |
| Dutch EEZ                    | Muddy deep bottom | Benthic Fauna | 3 indicators related to typological group sensitive to seafloor physical impact, based on (1) density, (2) biomass, (3) number of species | 1.1 (each) |
| Dutch EEZ                    | Muddy deep bottom | Benthic Fauna | 4 indicators related to typological group highly sensitive to seafloor physical impact, based on (1) density, (2) biomass, (3) number of species | 1.1 (each) |
| **BASQUE EEZ**               |         |                     | Cumulative contrib. of top-five indicators | 33                          |
| Offshore waters (> 200 m depth) | Benthic | Benthic habitat | Seabed affected by human activities | 9.9                        |
| Offshore waters (> 200 m depth) | Pelagic | Birds | Biological value Seabirds | 6.6                        |
| Offshore waters (> 200 m depth) | Pelagic | Mammals | Biological value Mammals | 6.6                        |
| Offshore waters (> 200 m depth) | Pelagic | Phytoplankton | Eutrophication indicator: Chlorophyll a, 90th percentile | 6.6                        |
| Offshore waters (> 200 m depth) | Sedimentary | Benthic Fauna | 2 indicators: (1) M-AMBI, (2) AMBI | 3.3 (each) |
| Offshore waters (> 200 m depth) | Sedimentary | Fish | Biological value Demersal Fish | 3.3                        |

(Continued)
### TABLE 3 | Continued

| Spatial Assessment Unit (SAU) | Habitat (general) | Ecosystem component | Indicator name | % contribution to assessment |
|-----------------------------|-------------------|---------------------|----------------|----------------------------|
| **PORTUGUESE CONTINENTAL SUBDIVISION** | | | **Cumulative contrib. of top-five indicators** | | |
| Continental_A2_600 | Marine | Birds | Biological Value Marine Birds | 17 |
| Continental_A2_600 | Pelagic | Zooplankton | Biological Value Zooplankton | 3.9 |
| Continental_B4_600 | Sedimentary | Benthic Fauna | Biological Value Benthic communities | 3.7 |
| Continental_B4_600 | Marine | Birds | Biological Value Marine Birds | 3.7 |
| Continental_B5_600 | Marine | Birds | Biological Value Marine Birds | 2.2 |
| **COASTAL BLACK SEA** | | | **Cumulative contrib. of top-five indicators** | | |
| Black Sea coastal | Pelagic | Fish | 4 indicators: (1) Mean length of Sprattus sprattus, (2) Catch/biomass ratio of S. sprattus, (3) Biomass of S. sprattus, (4) Sexually mature specimen of S. sprattus | 11.4 (each) |
| Black Sea coastal | Benthic | Fish | 4 indicators: (1) Mean length of Scophtalmus maximus, (2) Catch/biomass ratio of S. maximus, (3) Biomass of S. maximus, (4) Sexually mature specimen of S. maximus | 11.4 (each) |
| **SARONIKOS GULF** | | | **Cumulative contrib. of top-five indicators** | | |
| Saronikos Gulf | Marine | All Taxa | CIMPAL index (alien species) | 62 |
| Saronikos Gulf | Pelagic | Reptiles | % loss of spawning areas of sea turtle Caretta caretta | 4.4 |
| Saronikos Gulf | Pelagic | Fish | 4 indicators: Fishing mortality for (1) Engraulis encrasicolus, (2) Sardina pilchardus, (3) Merluccius merluccius, (4) Mullus barbatus | 4.4 (each) |
| Saronikos Gulf | Pelagic | Mammals | % Threatened mammals | 4.4 |
| Saronikos Gulf | Pelagic | Fish | % Threatened sharks | 4.4 |
| Saronikos Gulf | Pelagic | Fish | % of stocks that meet GES based on fishing mortality | 4.4 |
| Saronikos Gulf | Pelagic | Fish | % of stocks that meet GES based on reproductive capacity | 4.4 |
| Saronikos Gulf | Pelagic | Fish | % of stocks that meet GES based on reproductive capacity and biomass indices | 4.4 |
| **ADRIATIC SEA** | | | **Cumulative contrib. of top-five indicators** | | |
| Adriatic sea | Pelagic | Mammals | Tursiops truncatus, distributional range | 58 |
| Adriatic sea | Pelagic | Mammals | Stenella coeruleoalba, distributional range | 11.7 |
| Adriatic sea | Pelagic | Mammals | Grampus griseus, distributional range | 11.7 |
| Adriatic sea | Pelagic | Mammals | Ziphius cavirostris, distributional range | 11.7 |
| Adriatic sea | Pelagic | Mammals | Tursiops truncatus, abundance | 11.7 |

In case of equal contribution of several indicator values, all the indicator values are given. The contribution to the overall assessment (in %) of each indicator value is given. Numerical values are rounded. ES100, expected number of species in 100 individuals; AMBI, AZTI’s Marine Biotic Index; BQI, Benthic Quality Index; DKI, Danish Index; CIMPAL, Cumulative IMPacts of invasive ALien species; M-AMBI, multivariate AMBI; GES, good environmental status.

In NEAT, it is possible to weight the SAUs according to their perceived ecological relevance instead of their surface area; for example, biodiversity hotspots, important reproduction areas, marine protected areas, etc., could be given a higher weight than their area alone would imply. In this study, this option was not used in any of the case studies.

Uncertainty of the results is assessed based on Monte Carlo simulations, using the observed value as mean and the standard error value as the standard deviations, assuming a Gaussian distribution (Carstensen and Lindegarth, 2016). Based on these simulations, NEAT determines how often the sampled value falls into each of the five classes, and this distribution is reported. Therefore, the standard error values assigned to the indicators play a major role in the uncertainty associated with the final assessment result. This emphasizes the importance of careful evaluation of the standard deviation, particularly with indicators that have a high weight in the assessment.

### Evaluation of the Assessment Results

There are other tools to assess the status of marine systems, e.g., the Ocean Health Index (OHI; Halpern et al., 2012). This index has different concept and a much broader spatial scale, and a comparison between NEAT and OHI results (BD values presented in Table S6 in Selig et al., 2013) shows that the results are quite different (Table 5).
The OHI tends to give a more reduced range of status values (74–97) than those provided by NEAT (0.37–0.69) for these areas. The OHI does not provide a GES/non-GES status, but in general provides higher values than those by NEAT. The OHI study (Selig et al., 2013) has been applied globally, and includes a large variety of worldwide cases with great differences in setting and problems. In that context, e.g., the Mediterranean and the Baltic Sea seem to be in a (seemingly more homogenous) better state than e.g., waters around Africa or Indonesia and Philippines.

An interesting observation is that there is a negative rather than a positive correlation between these results, and those areas ranked low in NEAT (such as the Gulf of Finland and Kattegat) get high scores in OHI, while the best-scoring area in NEAT (Basque EEZ) gets lowest score in OHI (Table 5). This discrepancy is partly due to the fact that the OHI scores are given by country, thus covering larger areas than the case studies assessed here with NEAT. Therefore, the local status of a case study area may be masked by the results from the rest of the country in OHI. The NEAT results are reported here for the entirety of each of the case study areas, but where the case study area includes smaller SAUs, the results can be viewed for each of them separately as well, yielding even a more detailed geographical resolution.

Another factor possibly contributing to this discrepancy is the use of different indicators; the OHI assessment used publically available data with little local/regional detail, which can vary the final assessment when applying to regional scales (Halpern et al., 2014), while the current NEAT assessment used indicators specifically designed for marine status assessment. The species scores of OHI focused on the extinction risk of marine species (Selig et al., 2013), while the indicators in the NEAT assessments included a wider spectrum of indicators of species status. The OHI habitat scores were based on condition estimates of mangroves, coral reefs, seagrass beds, salt marshes, sea ice, and subtidal soft-bottom (Selig et al., 2013) while the NEAT assessments were tailored for each area.

The NEAT assessment results were in most cases in line with previous regional/local assessments, understanding, or known pressure gradients (Table 1, Figure 4). For example, The Baltic Sea biodiversity has been assessed by HELCOM (2009a, 2010) to be in poor to bad status in all of the three Baltic case study areas included in this analysis (Gulf of Finland, Lithuanian marine waters, Kattegat), being similar to the NEAT results but not to the OHI assessments. The difference between the NEAT and OHI results in these cases is probably largely due to eutrophication, which is documented to be major pressure threatening the ecosystem functioning of the Baltic Sea (HELCOM, 2009b, 2010). While it is reflected in the status of phytoplankton and water column habitats, and also affects the higher trophic levels of the food web (Österblom et al., 2007) and the seafloor (Karlson et al., 2002), it is not likely to be strongly reflected in the extinction threat of marine species (used in OHI), although it does affect the habitat scores, particularly seagrasses (Table S1 in Selig et al., 2013). Another factor affecting the discrepancy in the case of Finland is that the Gulf of Finland area has poorer biodiversity status than the Finnish marine waters on average (HELCOM, 2010).

### Table 4 | Number of indicator values per ecosystem component, and the relative weight of ecosystem components in each case study

| Ecosystem Component | Number of Indicator Values | Relative Weight (%) |
|---------------------|---------------------------|---------------------|
| Phytoplankton       | 4                         | 14                  |
| Zooplankton         | 3                         | 37                  |
| Benthic             | 9                         | 14                  |
| Fish                | 4                         | 6                   |
| Birds               | 1                         | 0                   |
| Mammals             | 0                         | 0                   |
| Reptiles            | 0                         | 0                   |
| Pelagic             | 0                         | 0                   |
| All                 | 28                        | 0                   |
| Water               | 14                        | 0                   |
| Benthic             | 25                        | 0                   |
| Phytoplankton       | 6                         | 0                   |
| Zooplankton         | 1                         | 0                   |
| Benthic             | 0                         | 0                   |
| Fish                | 0                         | 0                   |
| Birds               | 0                         | 0                   |
| Mammals             | 0                         | 0                   |
| Reptiles            | 0                         | 0                   |
| Pelagic             | 0                         | 0                   |
| All                 | 2                         | 0                   |
| Water               | 14                        | 0                   |
| Benthic             | 25                        | 0                   |
| Phytoplankton       | 6                         | 0                   |
| Zooplankton         | 1                         | 0                   |
| Benthic             | 0                         | 0                   |
| Fish                | 0                         | 0                   |
| Birds               | 0                         | 0                   |
| Mammals             | 0                         | 0                   |
| Reptiles            | 0                         | 0                   |
| Pelagic             | 0                         | 0                   |
| All                 | 2                         | 0                   |
| Water               | 14                        | 0                   |
| Benthic             | 25                        | 0                   |
| Phytoplankton       | 6                         | 0                   |
| Zooplankton         | 1                         | 0                   |
| Benthic             | 0                         | 0                   |
| Fish                | 0                         | 0                   |
| Birds               | 0                         | 0                   |
| Mammals             | 0                         | 0                   |
| Reptiles            | 0                         | 0                   |
| Pelagic             | 0                         | 0                   |
| All                 | 2                         | 0                   |
| Water               | 14                        | 0                   |
| Benthic             | 25                        | 0                   |
| Phytoplankton       | 6                         | 0                   |
| Zooplankton         | 1                         | 0                   |
| Benthic             | 0                         | 0                   |
| Fish                | 0                         | 0                   |
| Birds               | 0                         | 0                   |
| Mammals             | 0                         | 0                   |
| Reptiles            | 0                         | 0                   |
| Pelagic             | 0                         | 0                   |
| All                 | 2                         | 0                   |
| Water               | 14                        | 0                   |
| Benthic             | 25                        | 0                   |
| Phytoplankton       | 6                         | 0                   |
| Zooplankton         | 1                         | 0                   |
| Benthic             | 0                         | 0                   |
| Fish                | 0                         | 0                   |
| Birds               | 0                         | 0                   |
| Mammals             | 0                         | 0                   |
| Reptiles            | 0                         | 0                   |
| Pelagic             | 0                         | 0                   |
| All                 | 2                         | 0                   |
| Water               | 14                        | 0                   |
| Benthic             | 25                        | 0                   |
| Phytoplankton       | 6                         | 0                   |
| Zooplankton         | 1                         | 0                   |
In the North Sea, fishing is considered the main pressure, and the results show fish to be the ecosystem component in poorest status; the other assessed ecosystem components (birds, mammals, benthic fauna, and phytoplankton) were assessed to be in GES, with the exception of zooplankton that showed sub-GES (moderate) status (Figure 4). The Black Sea Coast case results obtained in this study also corresponded very well to known pressure gradients, such as nutrient enrichment affecting the status of the plankton community (Figure 4). Phytoplankton and benthic vegetation assessments correspond to category “poor” in the Varna Bay itself (Dencheva and Doncheva, 2014; Moncheva et al., 2015) as the most affected by anthropogenic pressure among the BSC sub-SAUs (Shtereva et al., 2012). The lowest benthic fauna score is also found there, which is fully in compliance with recently published results (National Report on the State and Protection of the Environment in Bulgaria, 2014). Similarly, the Basque area, which was previously assessed as being in good status, using a different methodology (Borja et al., 2011) also results in good status after applying NEAT; only mammals were assessed to be in sub-GES status (Figure 4).

In Saronikos Gulf the assessment results correspond to the ecological status categorization according to the WFD which is
poor in the sewage outfall area and moderate in the inner central gulf (Simboura et al., 2014, 2015, 2016). Aliens, fish including threatened sharks, and mammals contributed to the moderate status seen for the outer Saronikos and overall Saronikos. In general, the respective assessment results, although not definitive, are in line with pertinent studies (Frantzis, 2009; Katsanevakis et al., 2013; Papaconstantinou, 2014; Vasilakopoulos et al., 2015; Zenetos et al., 2015; Simboura et al., 2016) regarding the Greek marine waters. The Saronikos Gulf result obtained in this analysis was lower than the OHI assessment of the Greek waters, which was to be expected, as the Gulf is intensely exploited.

Results from the Norwegian part of the Barents sea indicated a general good status, which is in accordance with indicators of fish status on exploited large marine ecosystems (Kleisner et al., 2014; Coll et al., 2015), the report on the Barents Sea management plan (Sunnana et al., 2010) and the work from Certain et al. (2011). Nevertheless, several indicators indicated potentially degraded situations both in the coastal area and in the area of seasonal ice presence: (1) Along northern Norway coast, the current extent of kelp forest, an important component of fjords ecosystem and coastal landscape, cannot be considered as good in northern Norway. Kelp forests along the Norwegian and Russian coast were indeed dramatically grazed during the early 1970s and replaced by barren grounds dominated by sea urchins (Norderhaug and Christie, 2009). Though a progressive northward recovery of kelp forests extent is observed, its recovery status is still partial in northern Norway (Sivertsen, 2006; Rinde et al., 2014). (2) In northernmost part of the Barents sea, sea-ice extent is undergoing a particularly dramatic decrease (Parkinson et al., 1999) with a significant decrease rate of −3.5% per decade of winter ice extent (Sorteberg and Kvingedal, 2006) as a response to climate warming (Boitsov et al., 2014). This dramatic loss of habitat has consequences on the associated communities (Kovacs et al., 2011) as well as in the functioning of the Barents sea ecosystem as a whole (Wassmann et al., 2006). The growing evidence of impacts of climate change on this area rises the issue of exogenic unmanaged pressures on this system and the issue of shifting baselines for the definition of target values. In addition, there are still no indicators of the impact of trawling activities included in this assessment (see however Jørgensen et al., 2016).

For the Portuguese coast, the initial assessment officially provided in the scope of the MSFD (MAMAOI, 2012), presented a general environmental quality status higher than the NEAT results calculated in this study. This may be partly due to the fact that the present assessment did not include some special areas with a higher degree of protection (such as Berlangas' Marine Reserve and Professor Luiz Saldanha's Marine Park or Goringe Seafloor). These areas, which have restricted access by the public, are important for marine high trophic level species (e.g., marine birds, mammals), some of which were not included in the present assessment. Due to inconsistencies in the data (now being improved by projects such as MARPRO—Conservation of Marine Protected Species in mainland Portugal, http://marprolife.org), marine mammals, reptiles and benthic vegetation were not included in the current NEAT assessment, which may also contribute to the lower environmental quality results achieved by NEAT. The higher result reported by the OHI may be related to the methodology used for the scores' calculation, and may reflect more specifically the trend than the present environmental status.

An exception to the good correspondence between the current and previous assessments is the Adriatic Sea, where the assessment provided by NEAT appears too low considering the current trends, also reported in the scientific literature, and available information from expert opinions (Coll et al., 2010; Bastari et al., 2016). Despite the historical impacts on this shallow water basin, the Adriatic Sea is still characterized by a wide diversity of habitats, including rocky and soft bottoms, large estuaries and lagoons, seagrass meadows and in, its southern part, also deep-water environments. The habitat richness is reflected by a high biodiversity (Coll et al., 2012; Micheli et al., 2013b), with approximately 49% of the species described for the Mediterranean Sea (Boudouresque et al., 2009; UNEP, 2015) and a variety of endemic species (e.g., 18% of the endemic fish species of the Mediterranean; UNEP/MAP-RAC/SPA, 2015). Human activities and multiple stressors, and in particular bottom trawling, hydraulic dredging and habitat loss, are certainly still impacting the Adriatic Sea (Micheli et al., 2013a; Pusceddu et al., 2014). However, the overall environmental condition is not worsening with respect to the past decade. Eutrophication and dystrophic crises, related to the high nutrient discharge from the Po River combined with an alteration in water circulation, have caused hypoxia, anoxia and massive mucilage events, with consequent mortality of the benthic organisms, but the frequency of these events decreased significantly (or even disappeared) in the last decade (Degobbis et al., 2000; Danovaro et al., 2009). Thus, we hypothesize that the assessment of the environmental status obtained by using NEAT can be affected by the number and typology of data included in the specific exercise. An improvement of the number and type of the biological indicators (e.g., species or ecosystem functioning) could be crucial to obtain a more realistic classification of the marine environmental health of the Adriatic Sea.

Birds and mammals were found to be in poor status in many of the case study areas. This reflects the fact that seabirds are

| Country (Case study) | Status (NEAT) | Biodiversity score (OHI) |
|----------------------|--------------|-------------------------|
| Norway (Barents Sea) | 0.646        | 90                      |
| Finland (Finland)    | 0.401        | 97                      |
| Lithuania (Lithuania)| 0.583        | 89                      |
| Denmark (Kattegat)   | 0.394        | 93                      |
| Netherlands (North Sea) | 0.523    | 85                      |
| Spain (Basque Country) | 0.689     | 74                      |
| Portugal             | 0.566        | 83                      |
| Bulgaria (Bulgaria)  | 0.495        | 87                      |
| Greece (Saronikos Gulf) | 0.520   | 91                      |
| Italy/Croatia (Adriatic) | 0.370  | 86/89                   |

### TABLE 5 | Comparison of the biodiversity assessments obtained using the Ocean Health Index (OHI; data from Selig et al., 2013) and the Nested Environmental status Assessment Tool (NEAT) (this study) in the countries for which NEAT has case studies.
indeed considered as more threatened than any other comparable groups of bird species in general and display a faster trend of decline than other bird species during the last decades (Croxall et al., 2012). In addition, using IUCN Red list categories, it has been evidenced that, among seabirds, pelagic species of seabirds are disproportionately more threatened than coastal resident or coastal non-breeding visitor species (Croxall et al., 2012). Pelagic seabirds are particularly sensitive to disturbance as most species lay only a single egg, adults do not reproduce every year and usually reproduce several years after reaching sexual maturity (Furness and Camphuysen, 1997). Most seabird species display very large home range and thus integrate the state of the environment and impacts of pressures over larger scale.

The conservation status of marine mammals is of particular concern with an estimated proportion of threatened species ranging worldwide between 23 and 61% of species (Schipper et al., 2008). The North Atlantic region, which includes several of the cases studied here, is one of the areas where the proportion of threatened marine mammals is the highest, as shown by the low quality values in Barents Sea, Kattegat, and Basque case studies (Figure 4). The main reported threats explaining the bad status of marine mammals are a long history of harvesting, accidental mortalities through bycatch and collisions with vessel as well as a very large panel of pollutions (from sound pollution to contaminants and marine debris) and climate change (Schipper et al., 2008). The sensitivity of these species to changes in their environment might be related to their very slow population dynamics, low densities in correlation with their large body-size (Cardillo et al., 2008). Those life traits are also related with relatively large home range. As a consequence, indicators of marine mammals are usually measured over large scale, and they are difficult to monitor with precision, leading to higher uncertainty on many indicators (Taylor et al., 2007).

In two of the areas (Lithuania and Basque coasts), the indicator contributing the most to the final assessment was “the extent of the seabed significantly affected by human activities,” which is a direct indicator of pressure. This is interesting since some authors (e.g., Borja et al., 2013) have supported the use of pressures instead of assessing the environmental status, if there are not enough indicators. This should be done under the premise that if an area has no obvious pressures then any changes in the area must be due to natural changes which are outside the control of management and vice versa.

**Sensitivity Analysis**

The sensitivity analysis results show differences among the case studies in terms of how many indicator values are needed before the assessment results will show approximately the same results regardless of which indicator values are selected into the assessment (Figure 6). This implies that there is no universally sufficient number of indicator values needed to make a reliable assessment, but that the number varies among case studies. No clear patterns could be found among the 10 cases evaluated in this study that would indicate a number of indicator values of biodiversity components that can be considered sufficient regardless of the case study and its structure.

The variation in the assessment result depends on the set of indicator values that is available for the assessment. If the indicator values are close to each other, i.e., all indicating similar status, the variation in the results is naturally smaller. In contrast, if the different indicator values indicate very different status, e.g., some areas or biodiversity components are in good status while others are in bad, this naturally incurs a larger variation when a subset of these variables are selected, as e.g., in the Gulf of Finland.

These observations lead to the conclusions that if there is variation among the status of the geographical or biodiversity components in the study area, all of them should be covered by indicators if possible. Particularly the inclusion of high-leverage indicator values, i.e., those that have high weight and whose value differs from the overall mean, can change the assessment result. Therefore, the careful evaluation of the value and class limits of these indicators should be a priority.

**CONCLUSIONS**

The structured assessment forces us to critically evaluate the available indicator set in terms of ecological and spatial representativeness of each indicator. This framework highlights the gaps in the assessment as well as those parts that are well-represented by current monitoring and available indicators. This, in turn, helps in determining the best way to improve the quality of the assessment: (i) via developing additional indicators to fill in the gaps within the ecosystem approach (i.e., if not all the important trophic levels of key species/groups are covered in the existing indicator set), (ii) working to determine the optimal SAU for different categories of indicators that are targeted to assess various trophic levels and functions in the food web, as well as the HBT classification for each area, and (iii) working toward improving specificity, robustness, and pressure relevance of the indicators and enabling estimation of their standard errors.

The development of NEAT and this extensive testing with 10 case studies in very different European marine areas offers insight both to the status of the marine waters and to the state-of-the art of the available indicator assemblages as well as the development needs of the marine biological diversity assessment. The application of the tool will make the improvement and harmonization needs of the assessments visible and pave the way toward a harmonized assessment across large geographical scales.

In conclusion, we propose the following recommendations for the best practice in performing the environmental status assessment using NEAT:

- Careful attention needs to be paid particularly to the current status and class boundaries of the indicators that cover large geographical areas (such as mobile birds and mammals), as they tend to carry a lot of weight in the final assessment.
- In order to make the assessment comparable between the different sub-regions and areas in the regional seas and provide a harmonized assessment among the regional seas, the design of the assessment needs to be harmonized. Attention must be paid to the selection of ecosystem components, and definition
of size and hierarchy of the spatial assessment units as well as the definition of habitats.
- Consider the possibility of using different weighting for the individual indicator values, if that is ecologically more justified than using the weight based on the spatial area and habitat weighting.
- Contextualize the outputs on the basis of existing data. Different ecosystem components may present quite different data coverage, frequency, and data quality for the evaluation, and that may be reflected in the results. Consider carefully the standard deviation assigned to the indicators, but also consider how well the available indicators represent the ecosystem component and/or area as a whole.
- Consider not only the overall assessment, but the partial assessments (e.g., biological components or MSFD descriptors), as partial assessments can contribute to increased understanding of results and defining management measures for specific issues or areas.

**AUTHOR CONTRIBUTIONS**

LU, HB, and AB conceived the paper. The following partners provided the case studies: HB, SC (Barents Sea–Lofoten), LU, VK, HN (Gulf on Finland), GS (Lithuanian marine waters), CM, JA, JC (Kattegat), OB (Dutch North Sea), AB, MU (Basque Coast), JN (Portuguese continental subdivision), SM (Black Sea Coast), NP, MP, NS (Saronikos Gulf), RD, LC, AC, SB (Adriatic Sea). LU wrote the first draft. TB and CM contributed to the calculations and sensitivity analyses. LU, HB, AB, SC, AH, and MU contributed largely to the introduction and discussion. All authors contributed to the last draft and to the discussions.

**FUNDING**

This manuscript is a result of DEVOTES (DEvelopment Of innovative Tools for understanding marine biodiversity and assessing good Environmental Status) project, funded by the European Union under the 7th Framework Programme, “The Ocean of Tomorrow” Theme (grant agreement no. 308392), www.devotes-project.eu. MU is partially funded through the Spanish programme for Talent and Employability in R+D+I “Torres Quevedo.” Moreover, the monitoring of Saronikos Gulf was financed by the Athens Water Supply and Sewerage Company (EYDAP SA).

**ACKNOWLEDGMENTS**

The authors would like to thank DEVOTES Advisory Board members Paul Snelgrove and Simon Greenstreet for their constructive comments at the early phase of the work.

**SUPPLEMENTARY MATERIAL**

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fmars.2016.00159

**REFERENCES**

Andersen, J. H., Dahl, K., Göke, C., Hartvig, M., Murray, C., Rindorf, A., et al. (2014). Integrated assessment of marine biodiversity status using a prototype indicator-based assessment tool. *Front. Mar. Sci.* 1:55. doi: 10.3389/fmars.2014.00055

Andersen, J. H., Halpern, B. S., Korpipäne, S., Murray, C., and Reker, J. (2015). Baltic Sea biodiversity status vs. cumulative human pressures. *Estuar. Coast Shelf Sci.* 161, 88–92. doi: 10.1016/j.ecss.2015.05.002

Andersen, J. H., and Stock, A., Heinänen, S. M., Mannerla, and Vinther, M. (2013). Baltic Sea biodiversity status vs. cumulative human pressures. *Estuar. Coast Shelf Sci.* 161, 88–92. doi: 10.1016/j.ecss.2013.09.008

Andersen, J. H., and Stock, A., Heinänen, S. M., Mannerla, and Vinther, M. (2013). Baltic Sea biodiversity status vs. cumulative human pressures. *Estuar. Coast Shelf Sci.* 161, 88–92. doi: 10.1016/j.ecss.2013.09.008

Andersen, J. H., Halpern, B. S., Korpipäne, S., Murray, C., and Reker, J. (2015). Baltic Sea biodiversity status vs. cumulative human pressures. *Estuar. Coast Shelf Sci.* 161, 88–92. doi: 10.1016/j.ecss.2015.05.002

Artegiani, A., Bregant, D., Paschini, E., Pinardi, N., Raicich, F., and Russo, A. (1997a). The Adriatic Sea general circulation. Part I: air–sea interactions and water mass structure. *J. Phys. Oceanogr.* 27, 1492–1514.

Artegiani, A., Bregant, D., Paschini, E., Pinardi, N., Raicich, F., and Russo, A. (1997b). The Adriatic Sea general circulation. Part II: baroclinic circulation structure. *J. Phys. Oceanogr.* 1515–1532. [Epub ahead of print].

Bastari, A., Micheli, F., Ferretti, F., Pusceddu, A., and Cerrano, C. (2016). Large marine protected areas (LMPAs) in the Mediterranean Sea: the opportunity of the Adriatic Sea. *Mar. Policy* 68, 165–177. doi: 10.1016/j.marpol.2016.03.010

Beaumont, N. J., Austen, M. C., Atkins, J. P., Burdon, D., Degraer, S., Dentinho, T. P., et al. (2007). Identification, definition and quantification of goods and services provided by marine biodiversity: Implications for the ecosystem approach. *Mar. Pollut. Bull.* 54, 253–265. doi: 10.1016/j.marpolbul.2006.12.003

Berg, T., Murray, C., Carstensen, J., and Andersen, J.H. (2016). *Manual, Guidelines and Software for Biodiversity Assessment. DEVOTES Deliverable D6.3.* Available online at: http://www.devotes-project.eu/download/25439/

Birk, S., Bonne, W., Borja, A., Bru cet, S., Courrat, A., Poikane, S., et al. (2012). Three hundred ways to assess Europe's surface waters: an almost complete overview of biological methods to implement the water framework directive. *Ecol. Indic.* 18, 31–41. doi: 10.1016/j.ecolid.2011.10.009

Boitsov, V. D., Karsakov, A. L., and Trosimov, A. G. (2014). Atlantic water temperature and climate in the Barents Sea, 2000–2009. *ICES J. Mar. Sci.* 69, 833–840. doi: 10.1093/icesjms/fss075

Borja, A., and Collins, M. (2004). *Oceanography and Marine Environment of the Basque Country, Vol. 70.* Elsevier Oceanography Series, 616.

Borja, A., and Rodriguez, J. G. (2010). Problems associated with the ‘one-out, all-out’ principle, when using multiple ecosystem components in assessing the ecological status of marine waters. *Mar. Pollut. Bull.* 60, 1143–1146. doi: 10.1016/j.marpolbul.2010.06.026

Borja, A., Dauet, D. M., and Grémaire, A. (2012). The importance of setting targets and reference conditions in assessing marine ecosystem quality. *Ecol. Indic.* 12, 1–7. doi: 10.1016/j.ecolid.2011.06.018

Borja, A., Elliott, M., Andersen, J. H., Berg, T., Carstensen, J., Halpern, S., et al. (2016). Overview of integrative assessment of marine systems: the ecosystem approach in practice. *Front. Mar. Sci.* 3:20. doi: 10.3389/fmars.2016.00020

Borja, A., Elliott, M., Andersen, J. H., Cardoso, A. C., Carstensen, J., Ferreira, J. G., et al. (2013). Good Environmental Status of marine ecosystems: What is it and how do we know when we have attained it? *Mar. Pollut. Bull.* 76, 16–27. doi: 10.1016/j.marpolbul.2013.08.042

Borja, A., Galparsoro, I., Iriogoi, X., Iriondo, A., Menchaca, I., Muxika, I., et al. (2011). Implementation of the European marine strategy framework directive: a methodological approach for the assessment of environmental status, from the Basque Country (Bay of Biscay). *Mar. Pollut. Bull.* 62, 889–904. doi: 10.1016/j.marpolbul.2011.03.031

Borja, A., Prins, T., Simboura, N., Andersen, J. H., Berg, T., Marques, J. C., et al. (2014). Tales from a thousand and one ways to integrate marine ecosystem components when assessing the environmental status. *Front. Mar. Sci.* 1:72. doi: 10.3389/fmars.2014.00072
