Theme 1: Biodiversity

This is a deliverable from the BalticBOOST project that was coordinated by HELCOM and co-financed by the European Union in 2015-2016 as part of the programme DG ENV/MSFD Action Plans/2016.

WP 1.1 Deliverable 1: Development of a biodiversity assessment tool

Partners: [1]Finnish Environment Institute (SYKE),Helsinki, Finland, [2]NIVA-Denmark Water Research, Copenhagen, Denmark, [3]Estonian Marine Institute (EMI), University of Tartu, Tallinn, Estonia

Authors: Henrik Nygård[1], Ciarán Murray[2], Kaire Torn[3], Vivi Fleming-Lehtinen[1], Georg Martin[3], Jesper H. Andersen[2], Samuli Korpinen[1]

Table of Contents

1. Introduction 2
2. Development of a biodiversity assessment tool (BEAT 3.0) 2
   2.1 Preparatory work 2
   2.2 Application of indicators 4
   2.3 Assessment structure and weighting of indicators 8
   2.4 Testing the tool 11
3. Assessing confidence 14
   3.1 Temporal coverage of monitoring data 14
   3.2 Spatial representability of monitoring data 15
   3.3 Confidence of classification 15
   3.4 Methodological confidence 16
   3.5 Integration of indicator confidences and overall assessment confidence 17
4. Visualization of the results 17
5. BEAT 3.0 manual 19
   5.1 R-script 19
   5.2 Input tables 20
   5.3 Step-wise manual to run BEAT 3.0 21
   5.4 Biodiversity assessment workspace 21
6. Comparison between BEAT 3.0 and the revised COM DEC 22
References 25

1. Introduction

The BalticBOOST WP 1.1 has developed a tool for assessing the status of biodiversity in the Baltic Sea that will be used in the upcoming 'State of the Baltic Sea report' that is produced by the HELCOM HOLAS II project. The tool was developed against assessment needs arising from the EU Marine Strategy Framework Directive (MSFD), the Baltic Sea Action Plan (BSAP) and the HELCOM Monitoring and Assessment Strategy. The first version of the 'State of the Baltic Sea report' will be produced by June 2017 and will make use of the HELCOM core indicators to produce integrated assessments of biodiversity, eutrophication and hazardous substances. The biodiversity assessment will be based on the HELCOM core biodiversity indicators and thus, the tool has been developed to function based on the core indicator properties. The development of the biodiversity assessment tool has been guided through two HELCOM workshops with participation of experts from the HELCOM Contracting Parties, the HOLAS II core team and the State and Conservation Working Group.

This report presents the work and developments carried out in BalticBOOST WP1.1. As preparatory work (Section 2.1) two tasks were carried out: 1) a review of available assessment tools to identify best practices and MSFD compliant methods and 2) a review of the HELCOM core biodiversity indicators to identify the properties of the indicators to be used and what requirements they set for the assessment tool. Based on these two tasks the development of the tool started by solving how to account for the different approaches used in the indicators to define the threshold values that determines good environmental status (GES) as well as creating methods to normalize the indicators (Section 2.2) and setting up the assessment structure (Section 2.3). The approaches and assessment structure were evaluated through test case scenarios (Section 2.4). To evaluate the confidence of the biodiversity assessment, principles for a confidence assessment to be run in parallel to the biodiversity assessment were developed (Section 3). A separate tool to visualize the indicator and assessment results was developed to ensure transparency of the assessment (Section 4). The BEAT 3.0 biodiversity assessment tool itself is constructed as an R-script (Section 5). Section 5 also provides instructions on how to run BEAT 3.0.

The tool as presented in this report was approved for use in HOLAS II by HELCOM HOD 51-2016.

2. Development of a biodiversity assessment tool (BEAT 3.0)

2.1 Preparatory work

2.1.1 Review of tools for integrated assessment of biodiversity

As a starting point of BalticBOOST WP 1.1, the applicability of existing assessment tools for MSFD purposes was analysed (HELCOM BalticBOOST Biodiv WS 1-2016 Background Doc 2). The review showed that integrated assessments tools with formalized assessment rules are scarce. Previously conducted assessments have usually had a limited scope focusing on only one biodiversity element, whereas broad biodiversity assessment tools have only been used in the Baltic Sea.

The usefulness of different integration methods was also analysed. It was concluded that in a context where different types of data and variables are included (as is the case in the HOLAS II assessment), the most suitable approach is to use a hierarchical assessment structure, where the integration to higher assessment levels are done by averaging indicators (possibly weighted) used at lower assessment levels, possibly using a one-out-all-out procedure at higher hierarchical levels. The only identified assessment tools utilizing such nested and hierarchical levels were the BEAT, MARMONI and NEAT tools. BEAT was used in the thematic biodiversity assessment carried out for the first HOLAS assessment (HELCOM 2009), the MARMONI tool (Martin & Torn 2014) was developed in a EU Life+ project, whereas NEAT (Berg et al. 2016) has been developed in the EU FP7 project DEVOTES.

The BalticBOOST Biodiversity workshop (HELCOM BalticBOOST Biodiv 1-2016) discussed the results from the review and noted that the strength of the MARMONI tool is in identifying areas not reaching GES, whereas BEAT and NEAT are better suited to assess how far the status is from the threshold value. Further, the workshop supported that the developing HELCOM biodiversity assessment tool should have the same basic features as the BEAT and NEAT tools. In practice, this meant recommending modifying the NEAT tool to fit the purposes of the HELCOM biodiversity assessment.

The basic principles of the NEAT tool are:

- nested indicator based assessments,
- aggregations based on spatial units and habitat types,
- weighting of indicators possible,
- gives an estimate of the assessment uncertainty.

The NEAT tool is also flexible and can be customized. The current NEAT tool utilizes only quantitative indicators. Inputs are required for the observed value, standard error, threshold value, and range of indicator values (min-max). Spatial aggregations are done from smaller to larger spatial scales, but not the other way, implying that all information is used at the largest spatial scale, whereas indicators assessed at large spatial scale are left out at smaller spatial scales.

2.1.2 Review of HELCOM core indicator properties

The HELCOM core biodiversity indicators were reviewed and their properties summarized to define the technical requirements of the developing tool (Table 1). It was noted that different concepts for defining GES have been used in different indicators (boundary, trend, interval and conditional approaches). Even within an indicator, different approaches may have been applied in different areas depending on the underlying data. It was also noted that in many indicators several parameters are assessed. The indicators are assessed at the spatial level seen appropriate and with ecological relevance of the studied parameter, spanning from HELCOM spatial assessment level 1 to 4[1], depending on the ecosystem component and underlying data. Thus, BEAT 3.0 should be able to cope with this diverse set of indicators. At the moment, the NEAT tool only runs on indicators assessed in relation to a threshold value, so technical solutions needed to be developed to include also indicators with other threshold concepts in BEAT 3.0.

Table 1. Selected properties of the HELCOM core biodiversity indicators. For the full review of HELCOM core biodiversity indicators, see HELCOM BalticBOOST Biodiv WS 1-2016 Background Doc 3.

HELCOM Biodiversity Indicators	Proposed GES type	GES direction	Number of parameters	Spatial assessment unit	Ecosystem component
Abundance of coastal fish key functional groups	boundary, interval, trend	above, in between, below, increasing trend, declining trend, no trend	2 (piscivores, cyprinids)	Sub-basin coastal waters	Fish
Abundance of key coastal fish species	boundary, trend	above, below, increasing trend, decreasing trend, no trend	1-3 (perch, flounder, cod)	Sub-basin coastal waters	Fish
Abundance of salmon spawners and smolt	boundary	above	1	Sub-basin	Fish
Abundance of sea trout spawners and smolt	boundary	above	1	Sub-basin coastal waters	Fish
Abundance of waterbirds in the breeding season	boundary (interval)	above (in between)	2 (no of species, population size)	Baltic Sea (sub-basins when enough data accessible)	Birds

[1]Spatial assessment levels; level 1: entire Baltic Sea, level 2: sub-basins, level 3: sub-basins divided in coastal and open sea areas, level 4: coastal areas divided into water bodies or water types.

HELCOM Biodiversity Indicators	**Proposed GES type**	**GES direction**	**Number of parameters**	**Spatial assessment unit**	**Ecosystem component**
Abundance of waterbirds in the wintering season	boundary (interval)	above (in between)	2 (no of species, population size)	Baltic Sea	Birds
Distibution of Baltic seals	boundary	above	3 x 3 (grey seal ,ringed seal, harbour seal; breeding distribution, moulting distribution, area of occupancy)	Sub-basin	Mammals
Number of drowned mammals and waterbirds in fishing gears	trend	declining trend	1	Sub-basin	Mammals, Birds
Nutritional status of marine mammals	boundary	above	2 x 2 (exponentially growing population, population at carrying capacity; by-caught, hunted)	Sub-basin	Mammals
Population structure of long-lived macrozoobenthic species	interval or boundary?	in between	1	Sub-basin	Zoobenthos
Population trends and abundance of seals	trend, boundary	increasing trend, no trend, above	3 x 2 (grey seal, ringed seal, harbour seal; population growth, abundance)	Sub-basin	Mammals
Proportion of large fish in the community	boundary	above	1	Sub-basins open sea	Fish
Reproductive status of marine mammals	boundary	above	1-4 (grey seal, ringed seal, harbour seal, harbour porpoise)	Sub-basin	Mammals
State of the soft-bottom macrofauna communities	boundary	above	1	Sub-basin, coastal areas water body	Zoobenthos
Zooplankton mean size and total stock	multimetric (boundary)	above	2 (mean size, total stock)	Sub-basin	Zooplankton

2.2 Application of indicators

2.2.1 Accounting for different GES approaches

Monotonic indicators

Monotonic indicators showing better environmental status with increasing or decreasing indicator value (Figure 1 a and b) are already in the form be inserted into the tool. Thus, no adjustments to these indicators are needed for the normalization (see 2.3.2) for use in BEAT 3.0.

Unimodal indicators

Unimodal indicators that show good environmental status within an optimal range, i.e. have two threshold values (Figure 1 c), need to be treated differently than monotonic indicators. To evaluate the distance to the threshold value, the indicator value is in BEAT 3.0 compared to the threshold value lying closer.

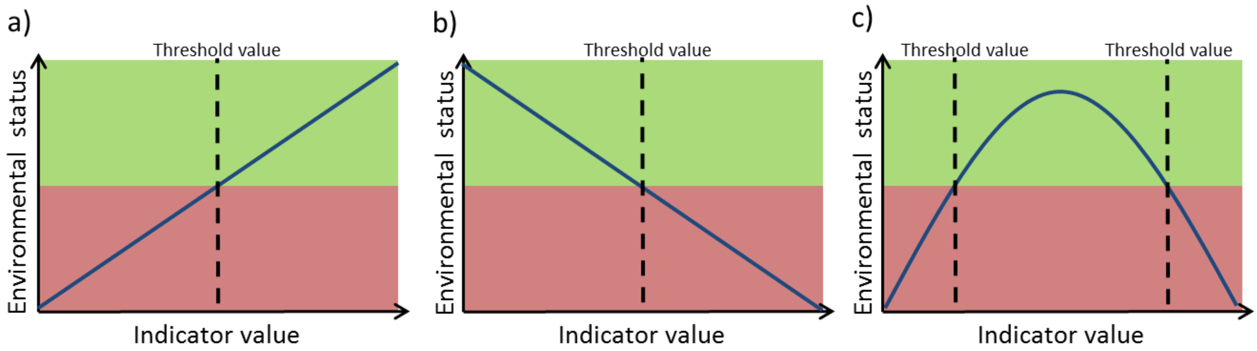


Figure 1. Different types of indicators and the setting of threshold values a) and b) monotonic indicators showing better environmental status with increasing indicator value, or better environmental status with decreasing indicator value, respectively, and c) optimal range of the indicator value.

Conditional indicators

Indicators, where several criteria need to be fulfilled e.g. GES is defined for several parameters and all need to be in GES in order for the indicator to be in GES, are referred to as conditional indicators. To keep a balanced structure in the assessment, i.e. not giving more weight to indicators that have several parameters, the tool needs to treat these indicators differently. In BEAT 3.0 this is solved in a way that all parameters are inserted in the tool, but when aggregating the results only the parameter showing the poorest result is included. This means that the indicators distance to the threshold value is decided by the parameter closest above the threshold value (if all parameters are in GES) or furthest below the threshold value (if any of the parameters are indicating bad status).

Trend indicators

Trend based GES definitions have been used for indicators where the amount of underlying data has been considered too low to set a fixed threshold value but the desired direction of change in the indicator value is known. Thus, for these indicators it is challenging to define the distance to the threshold value, if the maximum or minimum slopes of the trend cannot be defined. To include indicators with a trend based definition of GES in BEAT 3.0, two approaches are possible. 1) If the minimum and maximum slopes of the trend can be defined, they can be used to define the minimum and maximum values of the indicator (given that GES is defined as increasing/decreasing trend and threshold value thus 0). 2) If the minimum and maximum slopes of the trend cannot be defined, an approach based on expert judgment need to be applied. The general approach for the two-step expert judgment is shown in Figure 2. When presenting the indicator results in HOLAS II, trends are going to be used as defined in the HELCOM core indicator reports. The method presented in Figure 2 is only going to be used to define the input value to BEAT 3.0.

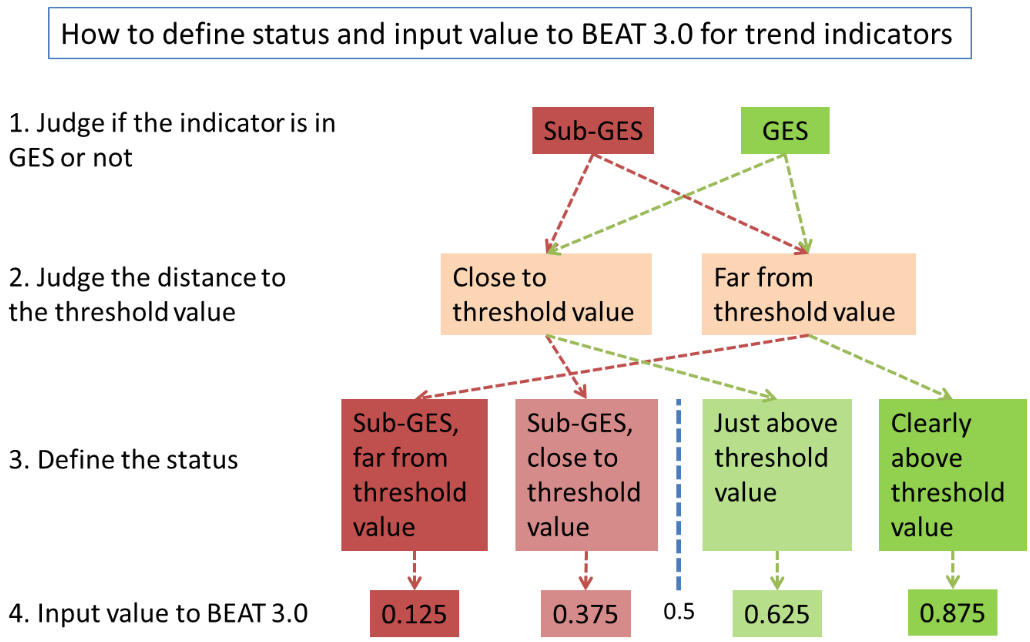


Figure 2. The general approach for how to include indicators with a trend based GES definition in BEAT 3.0.

Qualitative indicators

Qualitative indicators only assessing if the indicator is in GES or not in GES, without any indication of the distance to the threshold value get in BEAT 3.0 the input values 0.25 (sub-GES, mid-point of 0-0.5) or 0.75 (GES, mid-point of 0.5-1).

2.2.2 Normalization of indicators

The core indicators are based on very different types of data (units and scales) and thus, they need to be normalized in order to be included in a common assessment. Normalization is needed to estimate the distance of the indicator value to the threshold value (in addition to defining if the indicator is in GES or not) and to be able to compare this distance between indicators.

BEAT 3.0 normalizes and compares the distance to GES between indicators based on information on the full range of potential indicator values (minimum-maximum) and the threshold value. In BEAT 3.0, the normalized indicator value is called Biological Quality Ratio (BQR), and ranges from 0 ('bad status') to 1 ('high status') with the threshold value representing GES set at 0.6. Thus, when inserting the minimum and maximum values, carefulness is needed as the minimum and maximum values will have different meaning depending on the direction of change in environmental status compared to the indicator value. For example, when a decrease in the indicator value means an improved environmental status, the maximum indicator value reflects 'bad status'. If linearity between indicator value and environmental status can be assumed, defining minimum and maximum values in addition to the threshold value would be sufficient for normalization. If a more detailed classification of the indicator is available (e.g. WFD classes) this would add precision especially if linearity cannot be assumed. Defining minimum and maximum values of an indicator is straightforward when sufficient data covering the whole spectrum is available. This is however not often the case and due to the environmental degradation and limited data for some indicators the definition of both minimum and maximum values is challenging. Below, principles on how to set the minimum and maximum values, if not evident from the data, are presented.

Indicators with threshold value

The following approach is used to define minimum and maximum values for indicators where the whole range of possible values is not represented in the data. The approach requires that there is information on either the minimum or the maximum value:

Defining the maximum value

- if information on reference conditions in the sense of the WFD is available: reference condition = maximum value,
- other motivated basis: e.g. theoretically largest possible value.

Defining the minimum value:

- for time series covering deteriorated conditions: minimum value can be obtained from data,
- other motivated basis: e.g. theoretically smallest possible value.

The value that remains to be set (minimum or maximum) can be identified based on this information, if assuming linearity. The distance between the minimum value and the threshold value is then three units and the distance between GES and the maximum value is two units (see the provided example, Figure 3).

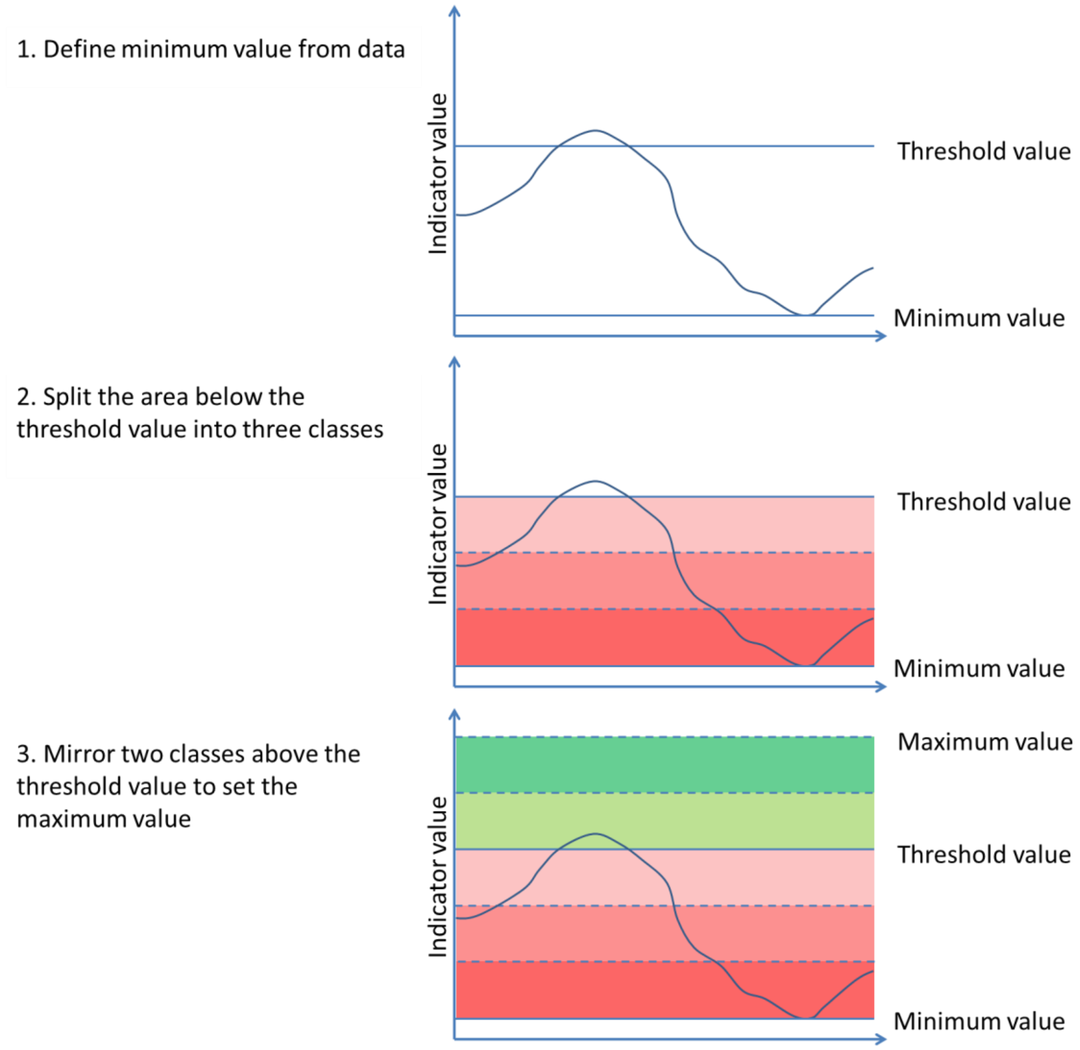


Figure 3. Example for an indicator with threshold value and data available for deteriorated conditions assuming linearity.

Indicators with optimal range

Unimodal indicators have two threshold values, meaning that the minimum and maximum values of the indicator will both reflect 'bad status'. In BEAT 3.0 this can be solved by providing two minimum and two maximum values, where the mid-point of the optimal range is defined as maximum value for the lower threshold value and minimum value for the higher threshold value.

Indicators with trend-based approach

For defining minimum and maximum values of indicators applying a trend-based approach, the threshold slope for a significant trend can be used as a proxy for a threshold value. This could be transformed into using degrees as the indicator "unit", so that the boundary slope is 0 degrees. The maximum value would be at +90 degrees ("above" the boundary, in the direction of improved state) and the minimum value at -90 degrees ("below" the boundary in the direction of worsened state). When the indicator value for trend indicators is determined based on the expert judgment approach (see figure 3), the minimum value is 0 and the maximum value is 1.

2.3 Assessment structure and weighting of indicators

BEAT 3.0 uses the ecosystem component structures as outlined in the draft revision of the Commission Decision on GES criteria for which the Marine Strategy Regulatory Committee gave a positive vote on 10 November 2016 (hereafter 'revised COM DEC'), i.e. five main ecosystem components (mammals, fish, birds, pelagic habitat and benthic habitat), divided into species groups/broad habitat types and further to species/habitats. The spatial structure follows the HELCOM Monitoring and Assessment Strategy, having four hierarchical levels as described in attachment 4 of the Strategy. The assessment is structured so that the relative contribution of ecosystem elements at each integration level is balanced, regardless of the number of underlying indicators. Indicators are then given weight based on the ecosystem element it is assigned to. Weights are only allocated to ecosystem elements and spatial assessment units represented by indicators. Thus, ecosystem elements and spatial assessment units not represented by indicators will not be included in the assessment. Area based weighing can be applied to aggregate the assessment results to higher assessment unit levels, but this will not be used in HOLAS II. Weighing based on ecological relevance of indicators is not considered, as recommended by HELCOM BalticBOOST Biodiv 1-2016, since it is challenging to scientifically and unequivocally justify. Figure 4 illustrates the assessment structure for ecosystem components and the distribution of weights to indicators. The spatial assessment structure is illustrated in figure 5.

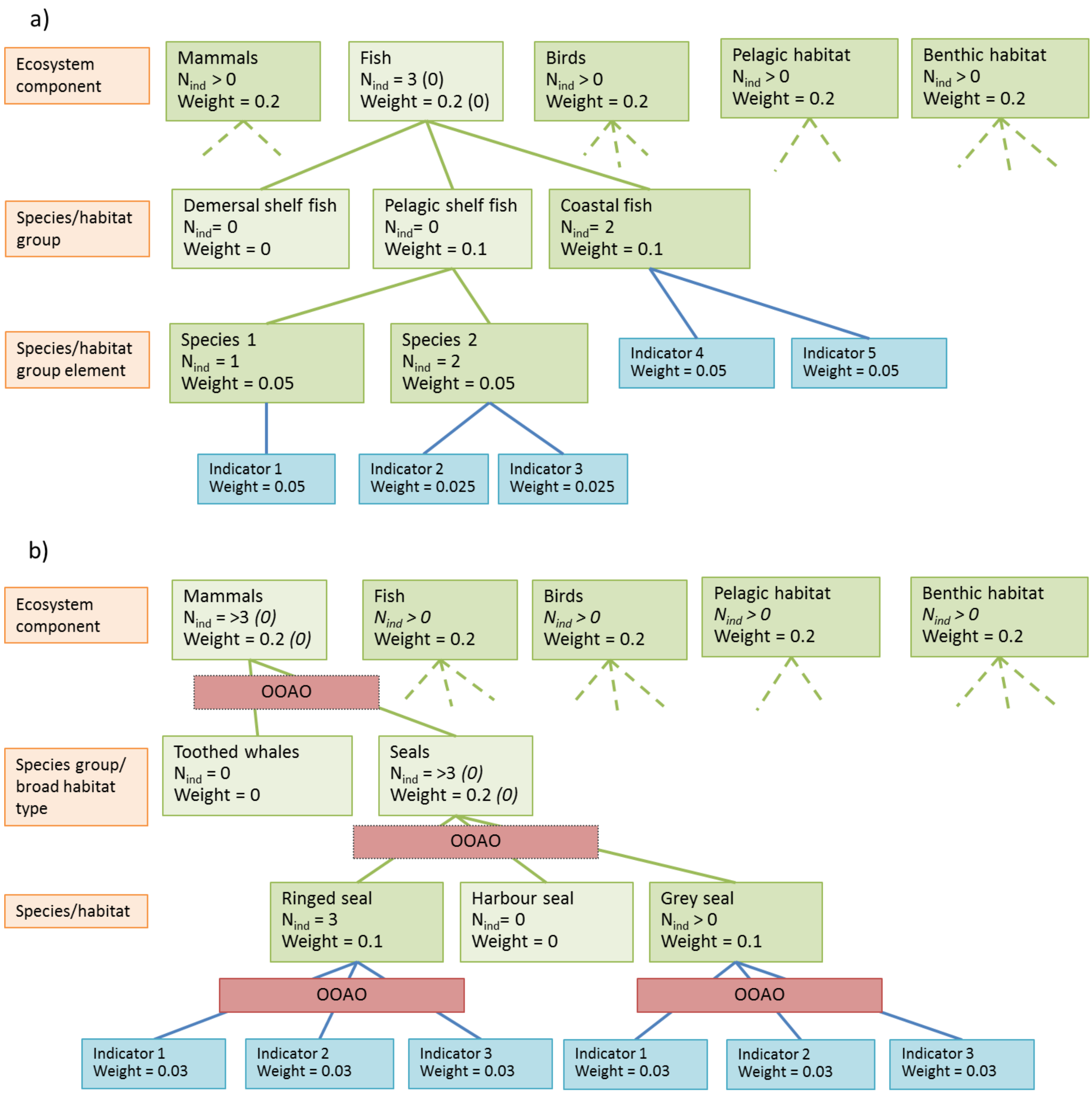


Figure 4. Assessment structure for ecosystem components and examples of distributing weights to indicators. a) An example on how weights are distributed to the indicators when indicators are assigned to different ecosystem element levels. For coastal fish indicators are assigned to the species group level, whereas for pelagic shelf fish indicators are assigned to species. b) For mammals the OOAO approach will be used in the integration.

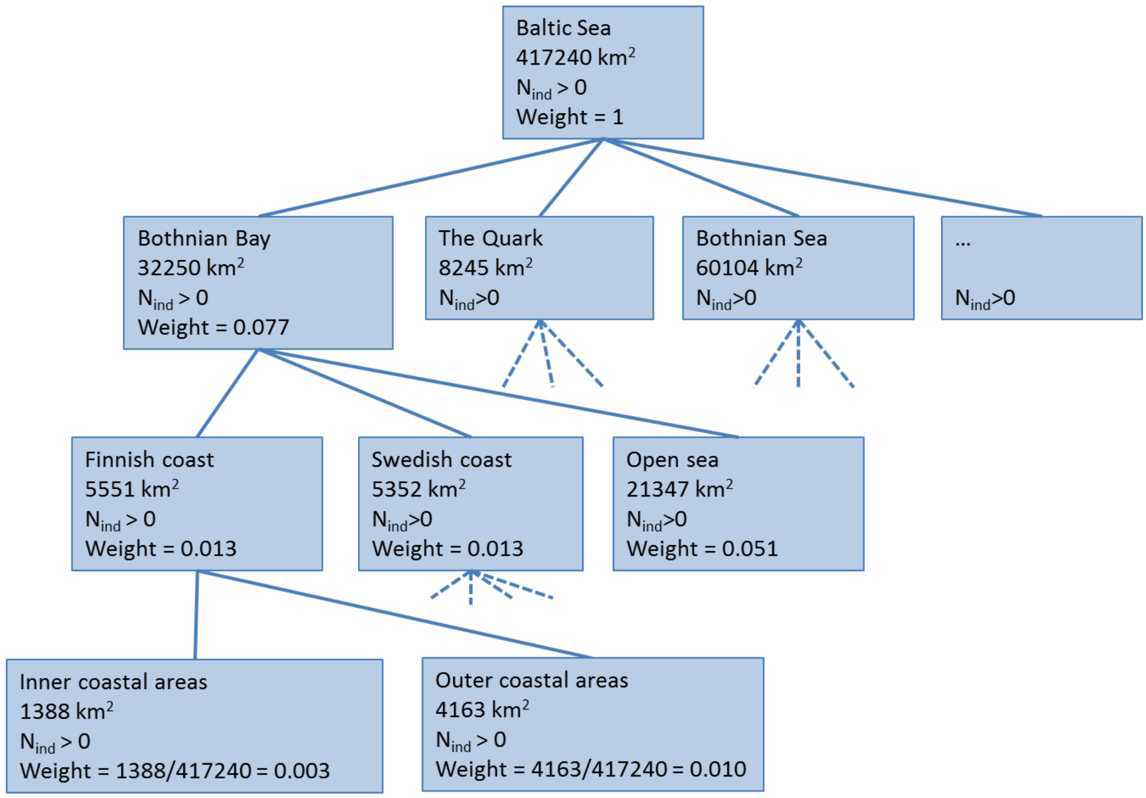


Figure 5. An example of the spatial assessment structure and the principle of assigning weight based on the area of the assessment unit. Note that this application will not be used in HOLAS II.

2.3.1 Aggregation methods for integrated assessment of biodiversity

Integrating indicators and ecosystem components

Using the weights assigned to the indicators, the tool starts from the lowest ecosystem component level and integrates indicators by weighted averaging. Stepping up to the following ecosystem element level, indicators assigned to this level are included. For mammals, integration following the OOAO principle is used to assess the status of species in accordance with the revised COM DEC which states that species covered by the Habitats Directive should be assessed by using the method provided in that Directive. Mammals also use OOAO for integration to species group and ecosystem component level as an interim solution as no guidance on this is given in the revised COM DEC.

Spatial scale and spatial aggregations

Results from an integrated assessment would be most useful at the most detailed spatial resolution. The present HELCOM core indicators, which will be used in the biodiversity assessment, are assessed at different spatial levels as defined in the indicator reports. This needs to be taken into account if integrated assessment results are wanted at the highest (most detailed) spatial resolution. A way to do this is to downscale the indicator results, i.e. give the same indicator result to all smaller assessment units situated in the larger assessment unit at which the indicator was assessed. Whereas aggregation methods for moving from a higher resolution to lower resolution could be motivated, using averaging approaches in order to obtain an overall assessment, moving form a lower to a higher spatial resolution may bring about some ecological issues. For example regarding the HELCOM core indicator on wintering birds, which is assessed on whole Baltic scale, it could become misleading to use the same indicator result on a finer scale in e.g. coastal areas of Bothnian

Bay which are normally ice-covered in winter. A possible solution would be to use weighting when down-scaling indicators, to give less weight in areas where the indicator becomes irrelevant. For HOLAS II purposes, the assessment will however, as agreed in HELCOM, be done at the spatial level defined as ecologically relevant for the indicators. This means that the results for the ecosystem components are presented at different spatial scales, tentatively birds at HELCOM assessment level 1, mammals at level 2, fish and pelagic habitats at level 3 and benthic habitats at level 4, without aggregating results across ecosystem components.

2.4 Testing the tool

2.4.1 Study areas and scenarios

The development version of BEAT 3.0 was tested in four case study areas: Gulf of Finland, Gulf of Riga, Bornholm Basin and Kattegat. In addition, an integrated test, based on the case study areas, was done for the Baltic Sea. The tests were carried out using available indicator data in the HELCOM biodiversity indicator reports and also utilizing the eutrophication indicators from the HELCOM EUTRO-OPER project. In addition, indicators developed in the MARMONI project were used in Gulf of Finland and Gulf of Riga to test how an increased number of indicators affect the assessment result.

Five different aspects regarding the biodiversity assessment were tested:

1) Assessment structure. Two different structures were tested: the 'criteria approach' and the 'species approach'. In the 'criteria approach', indicators are integrated to criteria and further to species group. In the 'species approach', indicators are evaluated per species before integrating to the species groups. These alternative structures were only tested for marine mammals, as indicators from other species groups did not support the species approach, i.e. indicator results were not available at species level.

2) Integration approaches. Weighted averaging at all steps or the one-out-all-out (OOAO) principle was used at different levels in the integration process. The alternative to end the assessment at the ecosystem component level and not have an overall assessment of biodiversity is also included in this set of tests.

3) Spatial representation. Three approaches were tested: i) using the spatial scales defined in the HELCOM indicator reports, ii) downscaling the indicators to HELCOM spatial assessment unit 4, and iii) same as in the second approach, but down-weighting indicators in areas where indicator experts considered the use of the indicator inappropriate, e.g. the species/habitat is not present in the sub-unit, and the indicator should thus not be assessed at this level. This information was collected from the indicator experts in a request for additional data.

4) Number of indicators. The test assessment was run for different sets of data using different numbers of indicators: using only HELCOM biodiversity core indicators, adding also relevant eutrophication core indicators (Secchi depth, oxygen, zoobenthos indices, macrophyte and phytoplankton indicators) and WFD indicators used nationally, as well as adding additional indicators suggested by the Contracting Parties. This scenario was complemented by including also indicators developed in the MARMONI project in the tests in Gulf of Finland and Gulf of Riga.

5) Assigning indicators to several criteria. The assessment result when using same indicators under several criteria was tested. The options were: i) indicators used only once, ii) same indicator used only once per descriptor, and iii) indicator used for all criteria it contributed to.

2.4.2 Summary of results and implications for the HOLAS II assessment

For testing the different scenarios a set of default choices (see Table 2) were made to reduce the amount of total combinations. Thus, we tested in total 14 different scenarios. In the tests, we used the criteria as described in the original MSFD Commission Decision (2010/477/EU). The final BEAT 3.0 version follows the revised COM DEC.

Table 2. The tested themes and alternative test scenarios. The marked scenarios were used as the default option when testing the other themes.

	Alternative test scenarios		
Themes for testing	1	2	3
Assessment structure	criteria based approach	species based approach	
Integration approaches	OOAO	weighted averaging	weighted averaging, OOAO at high level
Spatial representation	using indicator defined scales	down-scaling indicators to finest level	down-scaling indicators using weights
Number of indicators	BD core indicators	BD core + WFD and Eutro core indicators	all available indicators (including national indicators)
Indicators with multiple criteria	indicator used only once	indicator used once per relevant descriptor	indicator used for all relevant criteria

All the 14 test scenarios were successfully analyzed by the BEAT 3.0. Full results from the tests are presented in BalticBOOST Biodiv WS 2-2016 Doc 1. Here, a brief summary of the main points is presented.

Alternative structures

The two alternative scenarios of the integration structure, tested for mammals, did not differ in their result if the integration was made using a weighted averaging approach, as the weight of the indicators follow throughout the assessment. Thus, the way how to get to the ecosystem component through the structures, did not matter. However, if applying OOAO, the assessment result differs between the species and the criteria approach. Using the OOAO rule, the assessment outcome was showing poorer status when using the species based approach in Gulf of Finland and Gulf of Riga, whereas the criteria based approach showed poorer status in the Bornholm Basin and Kattegat. In the tests, the species based approach better reflected the status of species, whereas specific aspects of the ecosystem component, e.g. abundance, were better reflected in the criteria based approach. In a biodiversity assessment the species based approach gives a more easily communicable result.

Integration approaches

The different integration methods affected the assessment outcome. On Baltic Sea scale, the assessment results varied from 0.52 with weighted averaging, to 0.24 with OOAO at ecosystem component level, and further to 0.12 with OOAO at species group level. Integrating from criteria, averaging gave the result 0.60 and OOAO 0.45. Results from the case study areas showed the same pattern; applying OOAO resulted in poorer status. It is important to note that the integration method should not be chosen based on the result, but this is a fundamental principle of the assessment to be decided on. The test results showed that using OOAO for integration of indicator results will give high weight to single indicators. Weighted averaging can, however, fade out signals of concern. Using weighted averaging with OOAO at a high level will consider the indicators evenly, but still include the precaution in the final result. Ending the integration at the ecosystem component level is more informative than a single value for overall biodiversity.

Spatial representation

When testing the spatial representation of indicators only small differences (at highest 0.05 difference from the mean) in the assessment result was observed between the scenarios. However, when using indicators at their defined scales birds were not represented in any of the selected test areas, only at Baltic Sea level. If stopping the integration of indicators at the ecosystem component level 2 (i.e. not producing an overall biodiversity assessment), the spatial representation can be dealt with within the ecosystem components. The spatial representation of indicators within an ecosystem component is fairly similar, e.g. HELCOM bird indicators area assessed at the same spatial scale. However, if results are displayed at a finer spatial scale than the indicator is assessed on, downscaling the result will be needed. Based on the indicators used in the test scenarios it is recommended that indicators are downscaled in order to secure representability of all ecosystem components in sub-areas. The State & Conservation 5-2016 meeting noted that if the assessment is done separately for the ecosystem components, the spatial scale of the indicators is most often similar and no spatial scaling of the indicator would be needed. However, if different spatial scales are used for indicators within an ecosystem component the results should be presented at spatial level 3.

Number of indicators

When testing number of indicators, the overall biodiversity assessment score did surprisingly not vary much. The coverage of ecosystem components and criteria increased using more indicators and thus the overall confidence of the assessment is higher. Highest deviation from the mean (0.08) was observed in the Bornholm Basin. Within the ecosystem components the benthic habitat, and also pelagic habitat in Bornholm Basin and Kattegat, showed large variation between the scenarios due to the poor representability of indicators for these habitats in the biodiversity core indicators. The tests also showed that using only HELCOM core indicators for biodiversity will reduce the robustness of the assessment as the number of indicators per criterion and ecosystem component would be quite low. Including relevant eutrophication and WFD indicators (those that contribute to describing habitat condition) can improve the robustness, although all MSFD criteria cannot be assigned. However, based on the test results it is not possible to give scientifically based advice on how the choice of indicators should be made.

Use of indicators under several criteria

Using the same indicator under several criteria is not affecting the assessment result for the ecosystem components or overall biodiversity status, but changed the representability of indicators for the criteria. Using the same indicator once under several descriptors changed the descriptor results only slightly (<0.01 difference in assessment score), but included some more criteria and more ecosystem components per descriptor. Using the same indicator several times under the same descriptor changed the descriptor results a bit (0.07), but the differences within criteria and ecosystem component changed remarkably (up to 0.27). Using the species approach, the tool does not double count indicators if they are assigned to several criteria when using weighted averaging as integration method, but the results get influenced when using the criteria approach. Thus it is recommended that indicators are used only under one criterion per descriptor.

Evaluations and Implications for further tool development and use in HOLAS II

The second BalticBOOST Biodiversity workshop (HELCOM BalticBOOST Biodiv 2-2016), HOLAS II core team (HOLAS II 6-2016) and the State and Conservation working group (STATE & CONSERVATION 5-2016) evaluated the test results and recommended the use species approach for mammals as this is in line with the assessments done under the Habitats Directive. For other ecosystem components the criteria approach was favoured. Based on the revised COM DEC, the OOAO approach should be used for species covered by the Habitats Directive, which using HELCOM core indicators is currently limited to seal indicators. For other ecosystem components weighted averaging was favoured as integration approach. It was recommended that in HOLAS II integration of the ecosystem components to an overall biodiversity assessment should not be done. It was furthermore agreed in HELCOM to use the spatial resolution defined in indicator reports, as this is the ecologically most relevant scale. As indicators within an ecosystem component is most often assessed at the same spatial scale, no spatial aggregation would be needed in HOLAS II. In general, a recommendation to include as many indicators as possible to increase the confidence of the assessment was given, thus allowing the use of eutrophication indicators that can be used to assess habitat condition.

3. Assessing confidence

Confidence of the assessment result is in the NEAT tool presented as a probability distribution, based on information on the standard error of the indicator values entered into the tool. When it comes to the HELCOM core indicators, it was not be possible to apply a data-driven confidence evaluation based on the standard errors of the core indicators due to large differences in the indicators' underlying data and thus, an alternative approach needed to be applied. A classification system for evaluating confidence was thus developed. The confidence assessment is based on four categories (temporal, spatial, classification and methodological confidence) and informed by data to the extent possible, allowing the use of standard error to assess the confidence of classification. It is however noted, that a long term aim should be to develop a fully data-driven confidence assessment.

Confidence of the biodiversity assessment is based on the confidence evaluations of the indicators. The confidence need to be assessed separately for each indicator and assessment unit, and is inserted in BEAT 3.0 together with the indicator values. The confidence of all indicators is assessed in four categories: temporal coverage, spatial representation, confidence of classification and methodological confidence. For HOLAS II, indicator Lead and co-Lead country representatives assess the confidence in each category into the classes High, Intermediate or Low for the indicators used in the biodiversity assessment. In order to secure that confidence is assessed based on the same criteria for all indicators, the instructions on how to evaluate confidence were provided to the Lead and co-Lead country representatives that develops the HELCOM core indicators.

3.1 Temporal coverage of monitoring data

The aspect of temporal coverage of monitoring data considers the confidence of the indicator to include year-to-year variation in the indicator result. High confidence is considered to be achieved if monitoring data is available for all years in the HOLAS II assessment period (2011-2016), or for indicators where year-to year variation is not relevant, if the temporal monitoring requirements are met. Intermediate confidence is met when more than three years data is available from the assessment period and bad confidence is assigned to indicators with only one or two years of data during the HOLAS II assessment period (Table 3).

Table 3. Guidelines how to evaluate the temporal confidence of the indicator.

Score	**Evaluation: choose the score where the answer is 'YES' (to at least one question).**
HIGH	Does the monitoring data cover the entire HOLAS II assessment period? i.e. - if year-to-year variation occurs, are all years in the range 2011-2016 included? - if year-to-year variation does not occur, are the requirements for temporal frequency of monitoring met?
INTERMEDIATE	Does the monitoring data cover most of the HOLAS II assessment period? i.e. - if year-to-year variation occurs, are 3 or 4 years in the range 2011-2016 included?
LOW	Does the monitoring data cover the HOLAS II assessment period inadequately? i.e. - if year-to-year variation occurs, are only 1 or 2 years in the range 2011-2016 included?

	- if year-to-year variation does not occur, are the requirements for temporal frequency of monitoring not met? (Supplementary information: What is needed to improve)

3.2 Spatial representability of monitoring data

The spatial representability of monitoring data assesses how well the indicator covers the spatial variation within the assessment unit. Spatial representability is considered high if the data represents reliably at least 80% of the relevant habitat types occurring in the area, or in cases with a clear spatial gradient or patchiness in the parameter value, the monitoring is set to cover at least 80% of this variation. When the representability or variation (in case of gradients) is covered by 60-79% intermediate confidence is assigned to the indicator. Confidence is considered to be bad if less than 60% of relevant habitats or less than 60% of the variation in gradients are covered (Table 4).

Table 4. Guidelines for how to evaluate spatial confidence of the indicators.

Score	Evaluation: choose the score where the answer is 'YES' (to at least one question).
HIGH	Is the monitoring data considered to cover the full spatial variation of the indicator parameter in the assessment area? i.e. - does the data represent reliably at least 90% of the relevant habitat type(s) in the assessment area? - if a clear gradient or patchiness is shown in the parameter value, is the monitoring set to cover at least 90% of this variation?
INTERMEDIATE	Is the monitoring data considered to cover most of the spatial variation of the indicator parameter in the assessment area? i.e. - does the data represent reliably at least 70-89% of the relevant habitat type(s) in the assessment area? - if a clear gradient or patchiness is shown in the parameter value, is the monitoring set to cover 70-89% of this variation?
LOW	Is the monitoring data considered not to cover the spatial variation of the indicator parameter properly in the assessment area? i.e. - does the data represent reliably less than 70% of the relevant habitat type(s) in the assessment area? - if a clear gradient or high patchiness is shown in the parameter value, is the monitoring set to less than 70% of this variation?

3.3 Confidence of classification

The accuracy of the indicator result is primarily assigned as the standard error. If the standard error is not available a categorical approach is used. This is a compliance check by expert judgement of the probability that the indicator result clearly reflects that GES is achieved/not achieved. High confidence is assigned if GES has 'most likely' been / has not been achieved (by at least 90% probability). Intermediate confidence is judged if the probability is 'likely' (70-89% probability) and low confidence is judged if the probability of correctly indicating the status evaluation of the indicator is 'unsure' (less than 70% probability) (Table 5).

Table 5. Guidelines how to evaluate confidence of classification for the indicators.

For indicators that allow calculation of standard error for assessment evaluation:

Please provide the standard error of the evaluation data. The assessment tool will be able to score accuracy by combining this information with the status scoring.

For indicators that do not allow calculation of standard error for assessment evaluation:

Score	Evaluation: choose the score where the answer is 'YES'
HIGH	Does a compliance check to the threshold value show a clear signal whether GES has been achieved or not? i.e. - GES has been / has not been achieved by at least 90% probability
INTERMEDIATE	Does a compliance check to the threshold value show that values are generally clearly GES/sub-GES, though some outliers and variation in the data are present? i.e. - GES has been / has not been achieved by 70 – 89% probability
LOW	Does a compliance check to the threshold value not show clearly whether the data points are GES/sub-GES, and/or the overall evaluation is very close to the boundary? i.e. - GES has been / has not been achieved by less than 70% probability

3.4 Methodological confidence

The aspect of methodological confidence considers the quality of the monitoring methodology. High confidence is assigned if the monitoring has been conducted according to HELCOM guidelines (for parameters where these are available) and the data is quality assured according to HELCOM or other internationally accepted guidelines. Intermediate confidence is assigned if the monitoring has been conducted only partly according to HELCOM guidelines and/or the data originates from mixed sources, and is partly quality assured according to HELCOM or other international standards and/or the data is quality assured, but according to local standards. If monitoring has not been conducted according to HELCOM guidelines or the data has not been quality assured, the methodological confidence is considered bad (Table 6).

Table 6. Guidelines how to assess methodological confidence for the indicators.

Score	Evaluation: choose the score where the answer is 'YES' (to at least one question).
HIGH	For indicator parameters that have HELCOM guidelines for monitoring: has the monitoring been conducted according to these? Is the data quality assured according to HELCOM or other internationally accepted guidelines?
INTERMEDIATE	For indicator parameters that have HELCOM guidelines for monitoring: has the monitoring been conducted only partly according to these? Is the data from mixed sources, partly quality assured according to HELCOM or other international standards? Is the data quality assured, but according to local standards?
LOW	For indicator parameters that have HELCOM guidelines for monitoring: has the monitoring data not been collected according to these? Is the monitoring data not quality assured?

3.5 Integration of indicator confidences and overall assessment confidence

The confidence estimates are given in categorical form (low, intermediate and high confidence) and translated into numbers (0, 0.5 and 1) in the assessment tool. The translation to numeric form is essential in order to carry out the integration of the confidence through the tool. If standard error has been provided to assess the confidence of classification, BEAT 3.0 uses Monte Carlo simulations to estimate the probability of the indicator to be correctly classified. This probability will then be used as the value for the category 'confidence of classification'. Assessment confidence is calculated by BEAT 3.0 hierarchically, in two steps. Indicator confidence is first estimated separately for each indicator by averaging the confidences of the four categories. The indicator confidences are then combined together according to the aggregation principles of the status assessment. The final confidence result is presented in categorical form (High >0.75 ≥ Intermediate ≥ 0.5 > Low).

In addition to assessing the confidence of indicators, it is important to also consider how well the indicators cover the ecosystem components and MSFD criteria when assessing the overall confidence of the biodiversity assessment. If criteria or ecosystem elements are lacking indicators, i.e. are not assessed, this needs to be reflected in the overall confidence assessment. In BEAT, this implies reducing the confidence by 25% when criteria or ecosystem elements are lacking in the assessment (see Table 7 for specifications). If the criteria listed in the Table 7 are not fulfilled a penalty is accordingly applied to the confidence result (i.e. multiplying by 0.75). For ecosystem components, penalties at lower levels within the ecosystem component (e.g. species groups) are also taken into account. When several penalties are applied, they are multiplied (e.g. 0.75 *0.75 * confidence result). Similar approaches are also applied in the CHASE (hazardous substances assessment tool) and HEAT (eutrophication assessment tool).

Table 7. Reductions to confidence in the case MSFD criteria or ecosystem elements are insufficiently covered.

Applies to	**Criteria**	**Confidence penalty**
Species groups (mammals)	All HD annex II species covered	-25 %
Species groups (mammals)	All primary criteria covered	-25 %
Species groups (fish, birds)	All primary criteria covered	-25 %
Broad habitat types	All primary criteria covered	-25 %
Ecosystem components	All primary criteria covered	-25 %
Ecosystem components	All species groups/broad habitat types covered	-25 %
Criteria for ecosystem component	All species groups/broad habitat types covered	-25 %

4. Visualization of the results

The web based "Biodiversity assessment visualization tool" was developed to visualize and summarise the results of the biodiversity assessment (http://www.sea.ee/helcombd). Visualization is based on the result output tables produced by BEAT 3.0. The information on assessment score, confidence and number of indicators is used for visualization.

The visualization tool allows displaying the results of the biodiversity assessment according to different categories by spatial assessment units. For each spatial unit assessment results by criteria, ecosystem components, or species/habitat group can be selected.

The tool allows choosing between two options:

1) Results are sorted by ecosystem components at relevant spatial scale

According to HELCOM agreement, only the results for ecosystem components assessed at the spatial level relevant for the ecosystem component are shown. It means that for level 1 (Baltic Sea) only bird indicators are shown, level 2 (HELCOM sub-basins) 2 only mammals, level 3 (HELCOM coastal area and open sea) fish and pelagic habitats, and level 4 (national water types or bodies) benthic habitats. Under this option the assessment result can be visualized by criteria or by species/habitat group.

2) Results are presented with aggregation over all spatial levels

BEAT 3.0 have the ability to aggregate the assessment results to higher spatial scales and therefore visualization of several ecosystem components at same level is possible. Under this option the assessment result can be visualized by criteria, by ecosystem component or by species/habitat group.

The assessment results are presented as an interactive petal chart (Figure 6). Inner petal size indicates the assessment score value (between 0 and 1). Assessment results being in GES (score ≥0.6) are indicated with green petals and sub-GES results (score <0.6) in red. The outer circle shows the confidence class: high - light green, intermediate - yellow, low - orange. Moving the mouse pointer over the chart allows seeing the details: visualization property (code of criteria, name of the ecosystem component or species/habitat group), value of the assessment score, value of confidence and number of indicators used. The results can also be presented and downloaded as a pdf file showing the results in the petal chart and in tabular format (Figure 7).

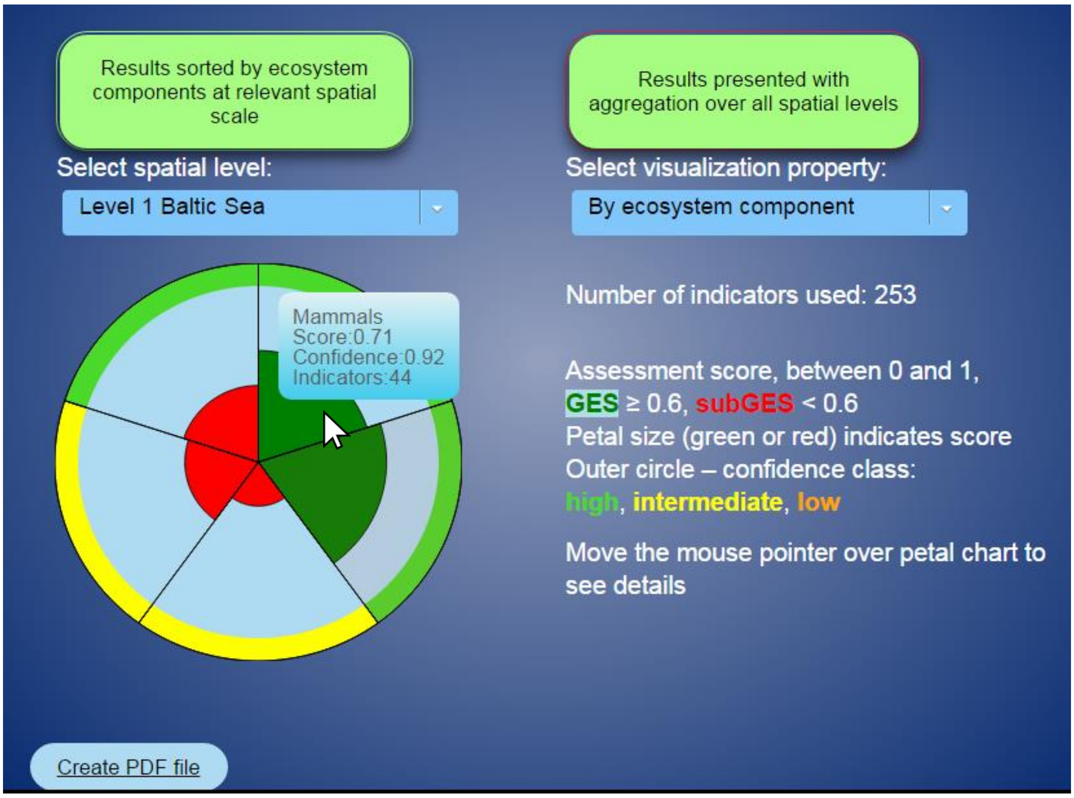


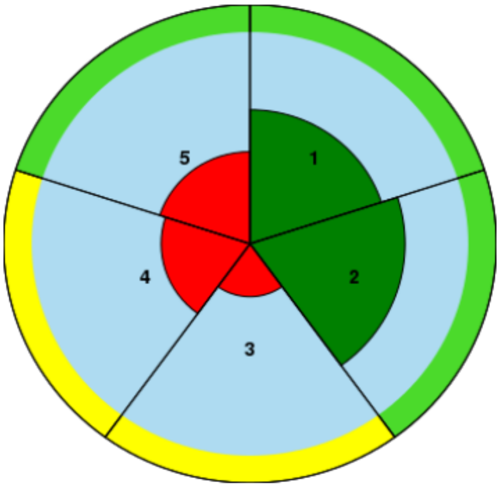
Figure 6. Example of visualization of assessment results.

Level 1 HELCOM sub-basin: Baltic Sea

Ecosystem component: Biodiversity

By ecosystem component

Evaluation is based on 253 indicators.
Assessment score in GES≥0.6, subGES<0.6.
Petal size indicates score on By ecosystem component.



		Ind. count	BQR	Confidence
1	Birds	2	0.63	high
2	Mammals	44	0.71	high
3	Fish	29	0.25	intermediate
4	Pelagic habitat	56	0.41	intermediate
5	Benthic habitat	122	0.44	high

Figure 7. Example of the pdf output file from the visualization tool showing the results as a petal chart and in tabular format.

5. BEAT 3.0 manual

5.1 R-script

BEAT 3.0 was coded as an R-script in order to provide a freely accessible and open tool. The script can be downloaded from: https://github.com/NIVA-Denmark/BalticBOOST.

The structure of the assessment is defined in input tables to the tool. The default spatial structure follows the HELCOM Monitoring and Assessment Strategy, whereas the ecosystem component structure follows the revised COM DEC. As a first step BEAT 3.0 reads the input tables, normalizes the indicators and assigns weight to them, as well as calculates the indicator confidence. The following step is, following the defined structures, integrating the assessment results to the different ecosystem component and spatial assessment unit levels. In the last step, the results are summarized and exported as output tables, separately for the biodiversity assessment and the confidence assessment.

5.2 Input tables

Input tables to the tool are:

Spatial assessment units (SAU.txt file) – a hierarchical list of the assessment units with four levels (according to the HELCOM spatial assessment unit levels 1-4). The area (km^2) of all spatial assessment units are specified here if applying the area-weighted spatial aggregation option.

Ecosystem components (EcosystemComponents.txt file) – a hierarchical list of the ecosystem components (birds, fish, mammals, pelagic habitats, benthic habitats) with four levels (1=Biodiversity, 2=Ecosystem component, 3=Species group/broad habitat type, 4=Species/habitat element). Each component is linked to the relevant higher level ecosystem component.

Descriptors (descriptors.txt file) – a list of the MSFD descriptors

Criteria (criteria.txt file) – a list of the MSFD criteria. This list is updated to follow the revised European Commission Decision on GES criteria.

Indicator catalogue (IndicatorCatalogue.txt file) – a list where the indicators are assigned to relevant ecosystem component and MSFD criterion.

Indicators (indicators.txt file) – table of observed value, minimum and maximum values, threshold value and confidence evaluations for the indicators linked to the spatial assessment unit. One row is added for each assessment unit the indicator is used in. Instructions on how to define minimum and maximum values for the indicators can be found in the request sent to the indicator Lead and co-Lead country representatives (HOLAS II 5-2016 Document 4-1 Annex 1).

Indicator group (ooao.txt file) – a list grouping indicators/parameters used in conditional indicators and indicators to be treated with the OOAO approach, i.e. using the parameter with poorest status classification in further integration steps.

Of these input tables, the spatial assessment units, ecosystem components, descriptors and criteria can be used as they are in the HELCOM and MSFD context, but if new indicators are added to the tool one needs to follow the steps and update the input files as outlined in the 'Step-wise description of the tool' (section 7).

5.3 Step-wise manual to run BEAT 3.0

1) Download the R script from https://github.com/NIVA-Denmark/BalticBOOST. Input files are found in the …/input folder.
2) Add the indicator to the indicator catalogue. Make sure to link the indicator with the correct ecosystem component ID and MSFD criteria. The ecosystem component ID is found in the EcosystemComponents.txt file
3) Insert the indicator results to the Indicators file. Add one row for each spatial unit the indicator has results for. If assessment is to be carried out on lower spatial assessment level than the indicator is assessed at, the information needs to be downscaled. This is done by adding the (same) indicator result to all relevant spatial assessment units at that level.
4) Specify the spatial assessment unit (ID is found in the SAU.txt file) and indicator ID (found in the IndicatorCatalogue.txt file).
5) Specify the indicator type (1: indicator value increasing/decreasing with improved/worsened environmental status, 2: indicator with an optimal range/interval).
6) Insert the minimum and maximum values of the indicator. Instructions on how to set the minimum and maximum values are found in HOLAS II 5-2016 Document 4-1 Annex 1. Make sure the minimum and maximum values are inserted correctly into the Bad and High columns, depending on if increasing value mean improved status (Minimum = Bad, Maximum =High) or decreasing value means improved status (Minimum = High, Maximum = Bad).
7) Insert the threshold value (ModGood column)
8) For type 2 indicators the optimal value is inserted in the High column. Minimum value is inserted in the Bad column, lower threshold value in the ModGood column, higher threshold value in ModGood2 column and maximum value in Bad2 column.
9) Insert the indicator result (obs column).
10) Insert the standard error of the indicator result (if available)
11) Define and insert the confidence scores (High = 1, Intermediate = 0.5, Low = 0) for each of the four categories: confidence of classification (ConfA), temporal coverage (ConfT), spatial representation (ConfS) and methodological confidence (ConfM). The confidence can be inserted in numerical or text format. Instructions on how to assess the confidence in the different categories are found in HOLAS 5-2016 Document 4-1 Annex 1. If standard error has been provided ConfA can be left empty.
12) If the indicator uses a conditional approach, i.e. several parameters with threshold values, all parameters and their results are inserted as separate indicators following the instructions above. The parameters are grouped in the ooao.txt file, where the indicator ID's of the parameters used in the indicator are given the same group ID.
13) Run the R script BOOSTbiodiv.R (make sure to specify the work directory location of the include.R file).
14) Result files are found in the …/results folder.

5.4 Biodiversity assessment workspace

To collect all information needed for the biodiversity assessment in HOLAS II a workspace was set up at HELCOM. In this workspace indicator results are collected, with the possibility for Contracting Parties to check and verify the results. For the biodiversity part, a test workspace was developed within the BOOST project where testindicator data were then collected in input tables that were used to test and develop by BEAT 3.0 and the visualization tool. The BEAT 3.0 result outputs are also stored in the workspace and utilized by the visualization tool. After finalisation of final indicator results and data, the test result tables will be replaced with real approved values and visualization tool will be made available via the HELCOM website displaying the assessment results.

6. Comparison between BEAT 3.0 and the revised COM DEC

Here a comparison between the final version of BEAT 3.0 and the revised COM DEC is presented. It should be noted that the development of BEAT 3.0 and the revision of the MSFD Commission Decision (2010/477/EU) were simultaneously taking place and BEAT 3.0 adjusted as far as possible according to the revised COM DEC. The majority of the HELCOM core indicators forming the base of HOLAS II were developed earlier and are thus not fully complying with the revised COM DEC.

Species

The revised COM DEC states that "*The status of each species shall be assessed individually, on the basis of the criteria selected for use, and these shall be used to express the extent to which good environmental status has been achieved for each species group for each area assessed*".

The relevant ecosystem components to consider for the Baltic Sea are mammals, birds, and fish. For mammals the relevant species groups are 'seals' and 'small toothed cetaceans'.

For birds, the original core indicators were based on a multispecies approach not differentiating the species groups. However, according to the latest indicator report the indicator can be differentiated to the following species groups: 'surface feeders', 'water column feeders', 'benthic feeders' and 'grazing feeders' and assessed per species group or species.

For fish, the commercially-exploited species are to be assessed according to D3 and used also under D1. Under D1 fish should be assessed according to the following species groups relevant to the Baltic Sea: 'coastal fish', 'pelagic shelf fish' and 'demersal shelf fish'. The assessment of fish can be based on the current core indicators according to species groups and the D3 ICES indicators are presented by species integrated to species groups.

Scale of assessment:

"Ecologically-relevant scales for each species group shall be used, as follows [including only selected information relevant for the Baltic Sea]:

- *for small toothed cetaceans, seals, birds, pelagic and demersal shelf fish: region or subdivisions for Baltic Sea*
- *for coastal fish: subdivision of region or subregion*
- *for commercially-exploited fish: as used under Descriptor 3.*

It could be noted that under the Habitat Directive countries in the Baltic Sea region have developed national approaches for assessing species under this directive, and that the approaches do not meet the 'scale of assessment' requirement in Commission Decision which states that for seals 'region of subdivision for Baltic Sea' should be assessed and furthermore that the 'scale of assessment' should be ecologically relevant. The HD approaches are generally developed to assess animals within the concerned countries borders, i.e. not on an ecologically relevant scale for the species. Thus, the regionally developed HELCOM core indicators for seals are well suited for meeting the requirements of the Commission Decision.

Use of Criteria:

The Commission Decision furthermore states that:

(a) *the assessments shall express the value(s) for each criterion used per species and whether these achieve the threshold values set;*

(b) *the overall status of species covered by Directive 92/43/EEC shall be derived using the method* *provided under that Directive. The overall status for commercially-exploited species shall be as assessed under Descriptor 3. For other species, the overall status shall be derived using a method agreed at Union level, taking into account regional or subregional specificities;*

(c) the overall status of the species group, using a method agreed at Union level, taking into account regional or subregional specificities.

These points will be covered in HOLAS II as follows:

a) The assessment of mammals will in HELCOM be made by assessing firstly the status of each core indicator per species. For seals, four indicators will be used in HOLAS II for grey seals while two indicators will be used for harbour seals and ringed seals. In the case of birds, two core indicators are used. At present they are assessing the ecosystem component, but results for species groups or species can be presented. As the GES definition presented in the bird indicator reports are not suitable to be used to assess species properly (baseline level periods would need to be specifies species-specific), it is proposed to assess birds per species groups. It needs to be noted that when assessing birds per species group the integration to 'Bird' level will be done by weighted averaging in the BEAT tool, i.e. not following the integration as described in the indicator report (GES = <25% of species deviating from the baseline). In the case of fish, the species group 'coastal fish' is assessed using key species and key functional groups. 'Pelagic shelf fish' and 'demersal shelf fish' are primarily assessed under D3 by species, but the results are also used for assessing D1.

b) Using the HD Directive method for overall status per species indicates the use of the OOAO approach between the criteria to arrive at the status per species. In HOLAS II this will only be relevant for mammals, as there are no core indicators for other species included in the HD annexes. BEAT will use the OOAO approach for mammals, whereas the results of indictors for fish and birds are weighted and averaged for the species/species groups.

c) In BEAT it is proposed that the overall status of the species group is based on weighted averaging. For species covered by the Habitat Directives the approach for assessing status of the species group is not defined in the 'GES Decision_v10.11.2016'.

Habitats

Habitats are according to the revised COM DEC to be assessed based on broad habitat types. Pelagic broad habitats have not been defined for the Baltic Sea, but the pelagic broad habitat types relevant to the Baltic Sea are only 'coastal' and 'shelf' (potentially also 'variable salinity'). 'Coastal' is not limited to the defined coastal areas in the WFD, but should be understood on the basis of physical, hydrological and ecological parameters. The benthic broad habitat types relevant in the Baltic Sea are the 'infralittoral' and 'circalittoral' habitat types. At present, there is a lack of indicators assessing the habitats for use in HOLAS II; only one indicator for pelagic habitats and only one indicator assessing soft bottom habitats.

Scale of assessment:
Subdivision of region or subregion, reflecting biogeographic differences in species composition of the broad habitat type.

The HELCOM assessment scale follows the division of sub-basins into coastal and open sea areas, where coastal follows the WFD definition and is further divided into water types/bodies. This subdivision can reflect the biogeographical differences in species composition of broad habitat types quite well, but a more precise analysis of differences in species composition within broad habitat types and HELCOM assessment sub-units is needed to define the appropriate subdivision.

Use of criteria: Pelagic habitats, D1C6:

The extent to which good environmental status has been achieved shall be expressed for each area assessed as:

a) an estimate of the proportion and extent of each habitat type assessed that has achieved the threshold value set;

b) a list of broad habitat types in the assessment area that were not assessed.

Units of measurement for the criteria:

D1C6: extent of habitat adversely affected in square kilometres (km2) per habitat type and as a proportion (percentage) of the total extent of the habitat type

Use of criteria: Benthic habitats, D6C4 and D6C5:

A single assessment per habitat type, using criteria D6C4 and D6C5, shall serve the purpose of assessments of both benthic habitats under Descriptor 1 and sea-floor integrity under Descriptor 6.

The extent to which good environmental status has been achieved shall be expressed for each area assessed as:

(a) for D6C4, an estimate of the proportion and extent of loss per habitat type and whether this has achieved the extent value set;

(b) for D6C5, an estimate of the proportion and extent of adverse effects, including the proportion lost from point (a), per habitat type and whether this has achieved the extent value set;

(c) overall status of the habitat type, using a method agreed at Union level based on points (a) and (b), and a list of broad habitat types in the assessment area that were not assessed.

Units of measurement for the criteria:

– D6C4: extent of habitat loss in square kilometres (km2) and as a proportion (percentage) of the total extent of the habitat type

– D6C5: extent of habitat adversely affected in square kilometres (km2) and as a proportion (percentage) of the total extent of the habitat type

For both pelagic and benthic indicators the HELCOM assessment is done at the assessment unit scale, i.e. one indicator value for the whole assessment unit. Currently, none of the indicators used in HOLAS II apply an area based approach with GES boundaries for proportion of adversely affected area. Thus, at present the proportion of adversely affected habitat type thus would need to be done using the areas of assessment units, i.e. proportion of assessment units adversely affected. For pelagic habitats, only one plankton-based indicator is anticipated to be used in the first version of HOLAS II, 'Zooplankton mean size and total stock'. For benthic habitats, only the *State of soft-bottom macrofauna community* indicator is anticipated to be operational with associated GES boundary for use in HOLAS II. However, this indicator is at present not suited for an area based assessment. D6C4 will not be covered in HOLAS II.

General observation on BEAT versus the revised COM DEC

- BEAT uses the same ecosystem component structure as defined in the revised COM DEC. HD requires an OOAO approach between the parameters assessed for each species.In BEAT this will only be applied for mammals, whereas birds and fish will apply weighted averaging. The method for overall status assessment of species groups and benthic broad habitats types is not set in the draft GES decision.
- BEAT utilizes the HELCOM assessment units, whereas the draft GES decision defines the assessment scales per species group and broad habitat types. These are not necessarily in contradiction to each other, but the need to reflect the proportion of adversely affected habitat types is not currently feasible as no indicator currently applies an area based approach.
- BEAT does not differentiate between primary and secondary criteria in the integration of indicators.
- 'Coastal' as used in the draft GES decision is not restricted to the WFD definition, but should take into account physical, hydrographical and ecological characteristics.
- BEAT allows integrated assessment results at higher ecosystem component levels (ecosystem components and biodiversity) as well as spatial scales (sub-basins, whole Baltic Sea) which is not required in the revised COM DEC, but can be useful in the HELCOM context. This possibility will, however, not be utilized in HOLAS II where the ecosystem components will be assessed separately at the relevant scale of each ecosystem component.

References

Berg T, Murray C, Carstensen J, Andersen JH (2016) NEAT – Nested Environmental Status Assessment Tool. Manual, version 1.2. (http://www.devotes-project.eu/wp-content/uploads/2016/07/NEAT-manual-v1.2.pdf)

HELCOM (2009) Biodiversity in the Baltic Sea – An integrated thematic assessment on biodiversity and nature conservation in the Baltic Sea. Balt. Sea Environ. Proc. No. 116B

Martin G, Torn K (2014) Integrated biodiversity assessment. In Aunins A & Martin G (eds) Biodiversity assessment of MARMONI project areas. MARMONI project report. p. 139-165 (http://marmoni.balticseaportal.net/wp/wp-content/uploads/2011/03/Biodiversity-assessment-report_24.03.2015.pdf)