Document title	Document 2 – Assessing confidence in the assessment
Category	CMNT
Submission date	30.8.2016
Submitted by	BalticBOOST

Background

This document presents how the BEAT 3.0 tool can estimate the confidence of the biodiversity assessment and what development steps the BalticBOOST project has taken in order to have suitable confidence information to the tool. The description of the BEAT 3.0 tool is presented in Document 1 of this workshop.

Action required

The workshop is invited to:

- evaluate the proposed approach of estimating assessment confidence,
- advise on how confidence should be assessed in the HOLAS II integrated biodiversity assessment

Assessing the confidence of the HELCOM HOLAS II biodiversity assessment

1. Introduction

The First HELCOM BalticBOOST workshop on the HOLAS II biodiversity assessment tool recommended that the integrated assessment of biodiversity will have a confidence assessment to express the underlying uncertainty in the assessment results. It was further recommended that this estimate will not influence the status assessment results but will be presented in separate.

In this document, the BalticBOOST project has taken forward the development of the confidence assessment and a proposal for the way forward is given.

2. Background to the BEAT 3.0 approach to calculate the assessment confidence

Assessment of confidence is calculated by BEAT 3.0 hierarchically, in two steps. Indicator confidence is first estimated separately for each indicator (see Section 3 below). The indicator confidences are then combined together according to the aggregation principles of the status assessment, taking into account the representability of the indicators (see Section 4 below).

This concept of calculating the confidence is based on the earlier HELCOM assessment tools HEAT, BEAT and CHASE (as well as HOLAS) and it is described in the respective HELCOM assessments (HELCOM 2009a[1], HELCOM 2009b[2],HELCOM 2010a[3], HELCOM 2010b[4], HELCOM 2014) as well as in Andersen et al. 2010[5], 2014[6], 2015[7]and 2016[8]. The concept includes not only the integration of indicator confidence but also other integration rules to evaluate the overall assessment confidence. For example, it is proposed that assessments of indicator groups that are based on single indicators are automatically given a low confidence (see section 4 below).

The confidence estimates in the abovementioned tools 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. A new opening in confidence estimates was taken in the DEVOTES project, where variance in the indicator values was used to present 'accuracy' of the indicator result. After Monte Carlo simulation, the accuracy was presented as the probability to be in the correct status class.

Based on HELCOM BalticBOOST Biodiv 1-2016 it was proposed that additional confidence criteria are included (in addition to the accuracy of the indicator result), and that these will take into account the spatial and temporal representability of the monitoring data as well as methodological confidence. The

[1] HELCOM (2009a) Eutrophication in the Baltic Sea—an integrated thematic assessment of eutrophication in the Baltic Sea region. Baltic Sea Environmental Proceedings No. 115B. Helsinki Commission, 148 pp

[2] HELCOM.(2009b).Biodiversity in the Baltic Sea – An integrated thematic assessment on biodiversity and nature conservation in the Baltic Sea, Baltic Sea Environment Proceedings 116B Helsinki Commission,188pp.

[3] HELCOM. (2010a). Hazardous substances in the Baltic Sea—an integrated thematic assessment of hazardous substances in the Baltic Sea. Baltic Sea Environment Proceedings, 120B, 1–116.

[4] HELCOM. (2010b). Ecosystem health of the Baltic Sea 2003-2007: HELCOM initial holistic assessment. Baltic Sea Environment Proceedings, 122, 1–63.

[5] Andersen, J. H., Murray, C., Kaartokallio, H., Axe, P., & Molvær, J. (2010). A simple method for confidence rating of eutrophication status assessments. Marine Pollution Bulletin, 60, 919–924.

[6] Andersen, J. H., Dahl, K., Göke, C., Hartvig, M., Murray, C., Rindorf, A., Skov, H., Vinther, M., Korpinen, S. (2014). Integrated assessment of marine biodiversity status using a prototype indicator-based assessment tool. Frontiers in Marine Science. http://journal.frontiersin.org/Journal/10. 3389/fmars.2014.00055/abstract

[7] Andersen J, Carstensen J, Conley DJ, Dropmh K, Fleming-Lehtinen V, Gustafsson BG, Josefson AB, Norkko A, Villnhäs A & Murray C. 2015. Long-term temporal and spatial trends in eutrophication status of the Baltic Sea. Biological Reviews, doi: 10.1111/brv.12221.

[8] Andersen J, Korpinen S, Murray C, Larsen MM, Green N, Høgåsen T, Gustavson K, Boalt E, Garnaga G, Haarich M, Mannio J & Strand J. 2016. A tool for integrated assessment of chemical status. Marine Monitoring and Assessment 188:115, DOI 10.1007/s10661-016-5121-x.

proposal is implemented in BEAT 3.0 which uses a categorical approach, which is informed by data, to assess the confidence. This is further discussed in Section 3 and 4 below.

3. Confidence information in the core indicators

The core indicators as they are presented in the fact sheets includes qualitative confidence information but in addition the BalticBOOST project asked the indicator experts to further define the confidence from four different aspects: temporal coverage, spatial representability, accuracy of estimate and methodological confidence. The indicator experts were asked to base their confidence evaluations on data, when possible. Replies were received for all core indicators, apart from those still under development. The four confidence aspects are explained one by one in the following paragraphs.

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 with no year-to year variation, the temporal monitoring requirements are met.
- Intermediate confidence is met when more than three years data is available from the assessment
- Bad confidence is assigned to indicators with only one or two years of data during the HOLAS II assessment period.

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 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.

3.3 Accuracy of indicator result

The accuracy of the indicator result is primarily assigned as the standard error. If the standard error is not available a categorical approach is carried out. This is a compliance check by expert judgement of the probability that the indicator signal 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).

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.

4. Integration of the confidence values

The confidence values in the indicators will be on the categorical scale (low, intermediate, high) which makes the integration of confidence cumbersome. The classes are transferred to the BEAT 3.0 as numbers (0, 0.5 and 1) in order to allow integration within the tool (Fleming-Lehtinen et al. 2015[9]). The confidence results are, however, expressed again as categories when the tool gives the results. All four values are fed into the tool together with the indicator result, to provide information for the next integrating step. They can also be used separately for possible additional analyses of confidence, and displayed in separate for purposes of transparency.

To evaluate the overall confidence of an integrated assessment, confidence information of single indicators is not sufficient. Information on how well the assessment covers the different aspects of biodiversity (e.g. species groups or criteria) also need to be taken into account in the assessment confidence. The practice of the earlier HELCOM assessments is recommended also for the BEAT 3.0 assessment, with the following implications:

- If any criterion lacks indicators, the overall confidence is automatically set as ‘low’.
- An indicator group (e.g birds, fish, mammals, pelagic and benthic habitat) with only a single indicator receives automatically ‘low’ confidence, regardless of the confidence of the indicator.

[9] Fleming-Lehtinen V, Andersen JH, Carstensen J, Łysiak-Pastuszak E, Murray C, Pyhälä M and Laamanen M. 2015. Recent developments in assessment methodology reveal that the Baltic Sea eutrophication problem is expanding. Ecological Indicators 48:380-388.